TRANSPOSON BASED MUTAGENESIS AND MAPPING OF TRANSPOSON INSERTION SITES WITHIN THE EHRLICHIA CHAFFEENSIS GENOME USING SEMI RANDOM TWO-STEP PCR

by

VIJAYA VARMA INDUKURI

B.S., ANDHRA UNIVERSITY, India, 2008

A THESIS

Submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Diagnostic Medicine/Pathobiology
College of Veterinary Medicine

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2013

Approved by:

Major Professor

Roman Reddy Ganta

Copyright

Abstract

Ehrlichia chaffeensis a tick transmitted Anaplasmataceae family pathogen responsible for human monocytic ehrlichiosis. Differential gene expression appears to be an important pathogen adaptation mechanism for its survival in dual hosts. One of the ways to test this hypothesis is by performing mutational analysis that aids in altering the expression of genes. Mutagenesis is also a useful tool to study the effects of a gene function in an organism. Focus of my research has been to prepare several modified Himar transposon mutagenesis constructs for their value in introducing mutations in *E. chaffeensis* genome. While the work is in progress, research team from our group used existing Himar transposon mutagenesis plasmids and was able to create mutations in *E. chaffeensis*. Multiple mutations were identified by Southern blot analysis. I redirected my research efforts towards mapping the genomic insertion sites by performing the semi-random two step PCR (ST-PCR) method, followed by DNA sequence analysis. In this method, the first PCR is performed with genomic DNA as the template with a primer specific to the insertion segment and the second primer containing an anchored degenerate sequence segment. The product from the first PCR is used in the second PCR with nested transposon insertion primer and a primer designed to bind to the known sequence portion of degenerate primer segment. This method aided in identifying the genomic locations of four *E. chaffeensis* mutants and also was valuable in confirming four other sites mapped previously by the rescue cloning method. This is the first mutational analysis study in the genome of an *Ehrlichia* species. Mapping the genomic transposon insertion sites is the first critical step needed for the continued

research to define the importance of the mutations in understanding the pathogenesis caused by the organism.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to express my sincere gratitude to my advisor Dr. Roman R. Ganta. I want to thank him for trusting in my abilities and giving me an opportunity to work as a graduate student under his guidance. His mentorship has played a pivotal role in the development of my scientific temperament. I have learned a lot from him and feel extremely fortunate to have had him as a mentor. I thank him for his timely help, feedbacks and patience throughout my project. One simply could not wish for a better or friendlier supervisor.

I would also like to show my sincere gratitude to all members of my advisory committee members: Drs. M. M. Chengappa and Melinda Wilkerson for their kind assistance and extensive evaluation of my Master's program. I can never forget the support Dr. Chengappa gave. I am very thankful to Dr. Wilkerson for offering the immunology and flow cytometry course.

I would especially like to thank Chuanmin Cheng for her guidance as well as friendship. Her capabilities helped me in every step of my project. Special thanks to my lab members: Huitao Liu, Arathy Nair, Tonia Ohlen, Chakravarthy Katragadda, and Gangadaar Thotakura.

I would like to thank my parents (Mr. Sita Rama Raju Indukuri and Mrs. Varalakhsmi Kumari Indukuri), my brother (Ajay Varma Indukuri), my sister-in-law (Satya Indukuri) and my grandparents (Mr. Padmanabharaju Bhupathiraju and Anasuya Bhupathiraju). I would like to thank all my friends. They made this journey enjoyable and a memorable experience.

This research project is funded by the grant from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, USA (AI070908).

Dedication

I dedicate this work to my family, friends and my major advisor Dr. Roman Reddy Ganta.

Chapter 1 - Literature review

Ehrlichia chaffeensis is an obligate intracellular Gram negative bacterium, and it is the causative agent of the disease human monocytic ehrlichiosis (HME). The organism belongs to the order Rickettsiales and the family Anaplasmataceae. The pathogen is transmitted by an infected tick *Amblyomma americanum*, commonly referred as the lone star tick [1]. HME is one of the most important tick-borne rickettsiale diseases in the USA. HME cases are reported from the south-central, southeastern, and Mid-Atlantic States of the USA, which corresponds to the geographical distribution of the lone star, tick. The symptoms of this disease may vary and may include persistent high fever, headache, malaise, and muscle aches [2, 3]. Ehrlichiosis can be fatal if not treated with an antibiotic, particularly in immune compromised and elderly patients [4]. The estimated case fatality rate (i.e. the proportion of persons who die as a result of their infection) is about 1.8%. Patients who are treated early may recover quickly on outpatient medication, while those who experience a more severe course of the disease may require intravenous antibiotics, prolonged hospitalization or intensive care. Diagnosis of HME is mainly based on clinical signs of the disease and followed by confirmatory laboratory test results. Some indicatives of HME in confirmatory laboratory tests include low white blood cell count, low platelet count and elevated liver enzymes, while the clinical signs resemble flue like illness with persistent high fever. HME is an emerging disease with number of cases increasing every year. The wide spread distribution of the reservoir host (white tailed deer) and the tick vector can also be a reason for the increased incidence of HME reported in recent years [5, 6].

Discovery

E. chaffeensis is an obligatory intracellular tick-transmitted rickettsiale bacterium in the family *Anaplasmataceae* of class *Alpha Proteobacteria.* It was first discovered in April 1986, in Arkansas in the peripheral blood smear of patients at Fort Chaffee, Arkansas [7]. The organism was isolated in cell culture and characterized by molecular techniques and was named as *E. chaffeensis* [8, 9]. It is closely related to the members of the genera *Anaplamsa*, *Wolbachia* and *Neorickettsia* belonging to the family *Anaplasmataceae.*

Statistics and Epidemiology

HME is identified as a reportable disease from 1999 and Centers for Disease Control and prevention (CDC) compiles the number of cases reported by state health departments. Compared to *E. chaffeensis*, number of cases with another species of *Ehrlichia* causing a disease in people, *E. ewingii,* are considerably less [10, 11]. The number of HME cases reported to CDC has increased steadily from 200 cases in the year 2000, to 961 cases in 2008 [10]. The incidence (the number of cases for every million persons) of ehrlichiosis increased similarly, from less than 1 case per million persons in 2000 to 3.4 cases per million persons in 2008 [12, 13]. During the same time period, the annual case fatality rate which means the proportion of ehrlichiosis patients that died as a result of their infections has declined [14].

Geographical distribution

HME is the most frequently reported disease from the southeastern and south-central United States. The areas from which cases are reported correspond with the geographic distribution of the lone star tick (*Amblyomma americanum*), which is associated with transmission of both *E. chaffeensis* and *E. ewingii.* Most cases reported to the CDC are from the states of Missouri, Oklahoma, Tennessee, Arkansas, and Maryland [15]. Mainly three states Oklahoma, Missouri, and Arkansas account for about 35% of all reported *E. chaffeensis* infections. Incidence of HME is also reported in China, Korea, Mali and Peru [13].

Morphology

Ehrlichia species are small, Gram-negative bacteria, round or ellipsoidal in shape. *E. chaffeensis* invade mononuclear phagocytes, such as monocytes and macrophages. The pathogen occupies cytoplasmic vacuoles, usually to form bacterial micro colonies known as morulae. *Ehrlichia* species cycle in nature between ticks and mammals, and can cause disease in many mammalian species. *Ehrlichia* species exist in two morphological forms during its life cycle, dense core cells which are the infective form and reticulate cells which are the non-infectious replicative form [16].

Pathogenicity

E. chaffeesnsis being an obligatory pathogen infects and resides in the monocytes and macrophages in the blood and other organs including spleen, liver, lungs and bone marrow of the host. It may also infect lymphocytes but it has a strong

tropism towards monocytes and macrophages. Usually the clinical signs develop in 1-4 weeks after the tick-bite [17]. Commonly the symptoms range from asymptomatic to mild flu-like symptoms, and sometimes might deteriorate into a fatal life threatening situation. Mostly common symptoms include malaise, fever, headache, muscle aches, chills, nausea and lymphadenopathy [10, 18]. Sometimes, there may also be cough, pharyngitis, swollen lymph nodes, vomiting, and possible changes in mental status[19]. Severe cases may include complicated situations such as septic shock-like syndrome, meningitis, organ damage [19]. Leucopenia, thrombocytopenia, and an increase in hepatic transaminase levels may be seen in the laboratory diagnosis [20]. Although people of all ages are susceptible to the disease, immune compromised people and elderly have greater chance of developing a severe disease [21].

Diagnosis and treatment

Since other bacterial and viral infections also have similar clinical signs, definitive diagnosis of HME requires the detection of pathogen by molecular methods. Thus, accurate description of tick-bite history and systematic laboratory diagnosis is very helpful in preliminary confirmation of this disease [18]. Routine laboratory diagnostic methods for preliminary examination include detecting peripheral blood smear, performing IFA and PCR. Confirmatory examination may also be possible by cell culture recovery of the pathogen [22-24]. For prevention and clearance of the infection, treatment with a tetracycline derivative is proven to be very effective [18, 25]. Specifically, doxycycline has better efficacy and is currently the drug of choice for treating HME patients. The drug binds to the 30S ribosomal subunit and prevents the

formation of peptide chains and therefore inhibits the bacterial protein synthesis and growth. There is no evidence for alternative treatments, as beta lactams, aminoglycosides and cephalosporins are not effective against *E. chaffeensis*.

Isolates

Culture system for growing *E. chaffeensis* aided in the recovery of various isolates of the organism recovered from human patients and from deer. The organism is cultured in various cell lines including canine histiocytoma cell line (DH82 cells), cells Human monocytic leukemia cells (THP 1), human cervical epithelioid carcinoma cells (HeLa cells), Vero cells, and murine fibroblasts [26]. Most of the isolates of *E. chaffeensis* are generated from HME patients [27]. A few isolates are also established from the bacteria recovered from the reservoir host, white tailed deer. The Arkansas isolate isolated is the oldest and the most widely used one in research [27].

Molecular Biology

The complete genome sequence of *E. chaffeensis* strain Arkansas is reported recently and is available at the NCBI database [28]. The *E. chaffeensis* genome is 1.18 Mb in size and consists of 1,115 open reading frames. It has a reductive genome and lacks genes for the biosynthesis of lipopolysaccharides and peptidoglycans. This is also observed in other related bacteria of the genera *Anaplasma* and *Wolbachia* belonging to the family *Anaplasmataceae* [29]. The GC % of *E. chaffeensis* is 20.1 and the average gene length is 840 base pairs. There are 604 genes coding for proteins with assigned functions and 314 are classified as genes coding for hypothetical proteins [30].

E. chaffeensis is unable to utilize glucose as a carbon or energy source and it is also auxotrophic for 14-17 amino acids [22, 30]. It must depend on the host for these amino acids and other metabolites. *E. chaffeensis* abundantly expresses more proteins involved in protein synthesis, and electron transport chain proteins [30]. *E. chaffeensis* has 40 genes encoding for transport and binding proteins. *E. chaffeensis* has the required genes for the biosynthesis of all necessary nucleotides, vitamins and cofactors like biotin, folate, FAD, NAD, CoA, thiamine and protohaem. The lack of required genes for the biosynthesis of lipopolysaccharide and common pili in *Ehrlichia* makes the envelope proteins important in providing a critical interface between the organism and its hosts. Whole genome sequence analysis revealed the identification of 49 genes that encode for envelope proteins in *E. chaffeensis* which corresponds to 6.6 % of the total encoded proteins [28].

Proteomics studies of *E. chaffeensis* identified one-fourth of the total open reading frames of the organism are expressed when it grows in macrophage and tick cell- cultures [31-33]. Mass spectrometric analysis identified 1,021 of *E. chaffeensis* proteins representing 92.3 % of the predicted bacterial proteomes [16] . Quantitative MS/MS analyses indicated that highly expressed proteins in *E. chaffeensis* included chaperones, enzymes involved in biosynthesis and metabolism, and outer membrane proteins such as P28/OMP-1 [16]. Proteomic analysis of infected host cells also showed that *E. chaffeensis* infection up regulated the expression of human proteins involved mostly in cytoskeleton components, vesicular trafficking, cell signaling, and energy metabolism, but down-regulated some pattern recognition receptors involved in innate immunity [34-36].

Outer membrane proteins

E. chaffeensis genome contains 22 paralougs of tandemly arranged *p28/omp-1* genes encoding immunodominant major outer membrane proteins [37-39]. These are clustered in a 29 kb locus which is present downstream of the transcriptional regulator gene *tr1*. This gene organization is conserved among other *Ehrlichia* species like [40-42]. The OMP1-P28 gene locus is among the most strain variable genomic regions [43]. Singu *et al* showed that these outer membrane proteins are differentially expressed in tick and macrophage cell lines [32].

Studies showed that immunization with a recombinant p28 protein protects mice from *E. chaffeensis* challenge. Polyclonal antibodies or monoclonal antibodies specific to p28 mediate protection of SCID mice from fatal infection with *E. chaffeensis*.

E. chaffeensis secretory system

Secretory systems in bacteria help in delivering or translocating effector molecules across the bacterial cells to their exterior. Secretory systems play a major role in the virulence and pathogenesis of a pathogenic bacterium. Intracellular pathogenic bacteria are known to translocate a wide variety of effector substrates into their host cells to help in manipulating the host cell environment for the bacterial survival. There are many types of secretory systems found in Gram negative bacteria, namely; type I secretion system, type II secretion system, type III secretion system (T3SS or TTSS), type IV secretion system (T4SS or TFSS), type V secretion system (T5SS) and type VI secretion system (T6SS).

Based on the genomic domain searches, *Hotopp et al* suggested that the bacterial type II, III, V and VI secretion components are not present in *E. chaffeensis [30].* T4SS and type I secretory system is identified in *E. chaffeensis* and the secreted products function as effectors in the host cells to facilitate the infection and intracellular replication of *Ehrlichia* [44, 45].

Cell Biology

E. chaffeensis infects the macrophages/monocytes, the cells with antimicrobial defenses. It is known that several pathogenic microorganisms have developed unique strategies in support of avoiding the activation of pattern recognition receptors. *E. chaffeensis* lost all the genes coding for the biosynthesis of lipopolysaccharide (LPS) and most of the genes for peptidoglycan biosynthesis [29]. Because of this, effective innate immune response targeted to the LPS of the organism is not possible. *E. chaffeensis* survives and replicates inside the midgut and salivary glands of its tick vectors. Arthropods also possess a strong innate immune system similar to vertebrates in support of overcoming the pathogen infections. It works by recognizing the pathogen-associated molecular pattern molecules (PAMPS) on bacteria. Some pathogenic bacteria also are known to escape the arthropod immune system as they do not express PAMPS [46]. The loss of genes coding the biosynthesis of LPS and peptidoglycan turns out to be an important adaptation strategy for *E. chaffeensis* and aids in evading vector immune system. [47, 48].

Lin *et al* have shown that *E. chaffeensis* requires cholesterol and have the ability of up taking the cholesterol from host [49]. With the aid of biochemical analysis and

fluorescence microscopy the authors demonstrated that *E. chaffeensis* contains high levels of membrane cholesterol [29]. Moreover, the absence of cholesterol causes ultra-structural changes to the organism's membrane structure. *E. chaffeensis* organisms when treated with cholesterol extraction agents, become incapable of infecting host cells and lose their viability [29]. Removal of cholesterol from the host cell in a culture system also impacts the bacterial ability to infect host cells. The dependency of *E. chaffeensis* on cholesterol for infection and replication *in vitro* suggests that increased blood cholesterol levels may increase the severity of disease in mammalian hosts.

Life cycle

E. chaffeensis requires both vertebrate and tick hosts to complete its lifecycle [16]. White tailed deer is the primary reservoir host. The pathogen is also detected in domestic dogs which may serve as a secondary host [50]. The lone star tick, *Amblyomma americanum* is the tick vector harboring the pathogen. Humans are considered incidental hosts acquiring infections from the bite of an infected tick bite [51-53]. After transmission to a mammalian host, *E. chaffeensis* replicates inside the monocytes/macrophages by subverting the immune responses of the host. The bacterium enters the host cells through endocytosis and resides in an intracellular compartment, in a phagosome [54, 55]. The inclusion does not acquire the components of late endosomes or lysosomes [56]. *E. chaffeensis* inclusions have the characteristics of early endosomes. They possess the early endosome markers like Rab 5, early endosome antigen 1 (EEA) and the vacuolar (H^+) ATPase. These inclusions fuse with endosomes containing transferrin and transferrin receptors [57].

E. chaffeensis forms dense intracellular micro colonies, termed as morulae, since they look like mulberries when blood smear of infected culture samples are stained and observed under a light microscope. Electron micrographs show that *E. chaffeensis* has enveloped inner and outer membranes [58]. The organisms are pleomorphic in shape and their size may vary from 0.2 to 2.0 µm in diameter. Inside the host cells *E. chaffeensis* appears in two forms, dense-cored cell and reticulate cells [58]. The two morphological structures are observed in both mammalian and tick cells [59-62].

Initially small dense cored cells enter the host cells during infection which are internalized and developed into large reticulate cells. The reticulate cells transform into dense cored cells prior to release [63, 64]. After a lag phase of growth which occurs for about 24 h, an exponential growth phase occurs which lasts for up to 72 h ,followed by a stationary phase from 72 h to 96 h [65, 66].

Host response to *E. chaffeensis* infection

Infection studies with *E. chaffeensis* were done in mice for studying the role of cellular and humoral immunity. Wild-type mice clear the infection within about 14 days [67]. On the contrary, the infection persists from several weeks to several months in mice with defective macrophage and T-cell functions [67] . Tlr 4 mutant mice produce decreased levels of nitric oxide and interleukin-6 showing the importance of macrophages in clearing infection. Mice without functional MHC-II genes fail to clear the infection which suggests that functional MHC system is essential for protection against *E. chaffeensis* [68, 69]. CD4+T-cell deficient mice are highly susceptible for *E. chaffeensis* infection whereas CD8+T-cell deficient mice are resistant to low doses of inoculum but susceptible to high doses [69].

Host cell gene expression during *E. chaffeensis* infection

E. chaffeensis infection alters the host gene expression. Zhang *et al* showed that the transcriptional levels of approximately 5% of host cell genes are significantly altered within 24 h following *E. chaffeensis* infection [62]. The genes with altered gene expression include those responsible for many important cellular processes, such as genes coding for apoptosis inhibition, cell cycle regulation and differentiation, signal transduction, proinflammatory cytokines, biosynthetic and metabolic proteins, and membrane trafficking proteins [62, 70] .

The microarray analysis of host cells during *E. chaffeensis* infection revealed that the bacterium manipulates genes related to three primary areas of the host responses. Upon infection the transcription of cytokines is repressed. Host cell cytokines like IL-12, IL-15, and IL-18 which are important for innate and adaptive immunity against intracellular bacteria are repressed. The host cytokines stimulates NK cells and T helper 1 cells to produce gamma interferon (IFN-γ), which is responsible for activating the macrophages to kill the phagocytized bacteria. IL-12 and IL-15 are known to activate NK cells and cytotoxic T lymphocytes, to kill the infected host cells. Repression of the host cytokines expression may be a survival adaptation mechanism of *E. chaffeensis* [62, 71].

Inhibition of apoptosis in infected host cells

Apoptosis is an innate immune mechanism of the host cells which restricts the replication of internalized pathogens. *E. chaffeensis* also upregulates NF-kB and other

apoptosis inhibitors. These help in enhancing the host cell survival. During infection host cells apoptosis inhibitors such as IER3 (immediately early response), BirC3 (baculoviral IAP repeat- containing protein), and BCL2 are induced. Apoptosis inducers such as BIK (BCL2 interacting killer), and BNIP3L (BCL2/adenovirus E1B 19-Kda interacting protein 3 like) are inhibited following infection with *E. chaffeensis*. Thus inhibitions of the host apoptotic mechanism provide the bacterium the chance for prolonged replication and survival [72, 73].

Manipulation of host membrane trafficking machinery

It is showed that *E. chaffeensis* inhibits the host cell genes involved in membrane trafficking. It inhibits the maturation of the early endosomes so that it can escape from the lysosome mediated destruction. During the first hour of infection *E. chaffeensis* represses the production of markers STX16 (syntaxin 16) and Rab5 [16]. The bacterium induces the production of vimentin a SNAP 33 reservoir and a component of membrane fusion machinery. *E. chaffeensis* may modulate phagosome-lysosome fusion by regulating the expression of Rab5 and SNAPs of the infected macrophage. Together, the detailed studies described in the literature demonstrate that *E. chaffeensis* is evolved to adapt to vertebrate and tick hosts and also alters host gene expression in support of its intracellular survival [16].

Mutagenesis in *E. chaffeensis*

One of the striking features of *E. chaffeensis* is its dual host life cycle and its prolonged persistence in both tick and vertebrate hosts. The mechanism of persistent is

not well studied. There may be several ways by which the pathogen could persist. One important hypothesis is that differential gene expression may be an important adaptive mechanism used by *E. chaffeensis* to support its continued survival in dual host environment [74]. Recent studies from our laboratory demonstrated that the pathogen's differential gene expression during growth in tick cells and mammals is a major contributor to its dual host adaptation [33, 75-77]. The importance of differential gene expression may be better understood by performing mutational analysis on the critical genes of *E. chaffeensis*. Creating mutations in *E. chaffeensis* remains a challenge because it is primarily an obligate intra-phagosomal pathogen with a limited survival in the extracellular environment [43].

Mutagenesis in bacteria is possible by creating insertion mutations at random or at specific genomic locations. Random mutagenesis approaches have been described in creating effective gene mutations in both Gram positive and Gram negative bacteria, including in *Rickettsial* organisms. For example, transposon based mutagenesis is described in *Rickettsia* and *Anaplasma* species [78]. Similarly, targeted mutagenesis is reported for Rickettsia species. Two methods of targeted mutagenesis include homologous recombination and Targetron based method [79]. Homologous recombination method is utilized for creating mutations in *Rickettsia* species [80]. Targetron mutation method is an efficient method in creating mutations in both Gram positive and Gram negative bacteria, but it is not described in rickettsiales. Targetron method involves the use of a modified mobile group 2 intron [81, 82]. Recently, studies have shown that Himar 1 transposase system is useful in creating insertion mutations at random in various bacterial genomes [81]. This method is also used in creating

mutations in intraphagosomal bacteria such as *A. phagocytophilum, A. marginale* and *Coxiella burnetti* Similarly other transposon systems, like TN5 and EZTN, have been used to create mutations in *C. burnetti* [83, 84].

Mapping of transposon insertion sites

One of the important challenges in random mutagenesis is the identification of genomic sites where mutations are introduced. Several molecular methods are described in the literature for mapping mutation sites. Total genomic DNA from transposed organisms can be assessed by Southern blot analysis targeting to a DNA inserted by the Transposon system. This method greatly aids in identifying inserts within a genome. Mutants can also be selected if inserts include an antibiotic resistance cassette. Specific insertion site identification within a genome requires additional molecular methods such as PCR and sequencing.

Methods described in the literature for mapping the random insertion mutations include rescue cloning, inverse PCR, nested PCR and semi random nested PCR [83]. One of the commonly used methods of identifying transposon insertion sites is a rescue cloning method. Rescue cloning is a commonly used method in which a genomic DNA containing the insertion sites was digested using a restriction enzyme that does not have restriction sites within the transposon insertion segments. Then the digested fragments are ligated into plasmid vector, transformed into an *E. coli* strain and the insertion-specific fragments are identified by screening for plasmids containing antibiotic resistance cassette originating from the insertion site [83]. Genomic sequences at the

insertion sites can then be identified by performing DNA sequence analysis of the recombinant plasmids by following standard molecular methods.

The semi random, two step PCR (ST-PCR)

Chun *et al* reported a variant of conventional PCR which aids in the rapid amplification of uncharacterized transposon tagged DNA sequences [85]. It is referred as the semi random, two step PCR (ST-PCR). This method involves two successive PCR reactions and two pairs of PCR primers. In the first reaction, one primer contains sequence complimentary to the transposon sequence and the second primer contains a 20-nucleotide sequence followed by a ten bases of degenerate sequence and a specific five nucleotide sequence GATAT. A subset of these primers anneal to an unknown DNA sequence near the transposable element. These primers will allow the initial amplification. The second primer pair is designed to be specific for the amplicons produced in the first round; one primer is targeted internal to the amplicons containing the transposon insertion sequence and the other primer is specific in binding to the 20-nucleotide sequence present in the semi random primer used in the first PCR. Insertion-specific sequences can be generated during the second PCR reaction when using the products of the first PCR as templates. The final products from the second round are gel isolated and subjected to DNA sequence analysis to identify the genomic insertion sites. ST-PCR is simple, efficient and rapid in identifying transposon insertions. Some software's applications are also available which can aid in predicting the location of transposon insertion sites. Several next-generation sequencing methods are also described in mapping transposon mutation sites [86].

SCOPE OF THE RESEARCH

Research has shown that *E. chaffeensis* employs differential gene expression for its survival in its two different hosts. Indeed, recent studies revealed that several proteins are differentially expressed in tick and mammalian host environments. Little is known about the molecular mechanism by which the bacterium changes its gene expression.

One of the best ways to study the impact of differential gene expression is to stop the gene expression of a subset of genes likely important and observe their effects on the bacterial growth in vertebrate and tick host environments *in vitro* and *in vivo*. This may be accomplished by creating mutations in bacterial genes. No successful mutagenesis methods are described in literature for *E. chaffeensis.* Our research group has been working for some time in evaluating different mutagenesis methods to create mutations in *E. chaffeensis*. Mutagenesis of a specific bacterial gene can be done in two ways; homologous recombination or by utilizing a modified group II intron-mediated mutational method. Although there is some success in cresting mutations, the targeted approach may not be ideal for *E. chaffeensis* because it is difficult to predict which genes are critical for the bacterial survival as most genes of the organism are expressed during its replication cycle in vertebrate and tick host cell environments. Indeed, research described by our group demonstrate that the targeted mutagenesis methods do aid in generating gene inactivation mutations, but the mutant organisms survive only for a very short span of time under *in vitro* culture conditions. Random mutagenesis by utilizing a transposon insertion mutation method may aid in identifying genomic targets essential for the *in vivo* growth while not impacting the organism's replication *in vitro*.

Mutagenesis experiments using Himar transposon system have been proven useful in several rickettsial pathogens. Our research team, in collaboration with Dr. Ulrike Munderloh's group, has also been successful in creating random mutations at various genomic locations in *E. chaffeensis*. The identification of the genomic sites where the random mutations are present is an important goal in furthering our knowledge about assessing the impact of mutations on the pathogen's growth. The primary focus on MS my research, therefore, is to fill-in this important gap of knowledge. In particular, the main focus of my research has been to map the mutation sites in *E. chaffeensis* genome.

Materials and Methods

Polymerase chain reaction (PCR)

All PCRs were set up in a final reaction volume of 25 µl. The reaction mixture consisted of a final concentration of 1X PCR reaction buffer containing 50 nm of MgCl2, 10 nm of dNTP's, 10 pico moles each of forward and reverse primers, about 1 ng of template DNA, and 1 unit of Taq DNA polymerase (Invitrogen technologies, Carlsbad, CA). The PCR temperature cycles followed an initial heating at 95°C for 2 min, followed by 40 cycles of 95°C denaturation for 30 sec, primer annealing for 30 seconds carried out at appropriate temperatures calculated for each primer set, 72°C extension for 30 sec. The extension temperature for platinum High fidelity DNA polymerase was 68°C. The extension temperatures were increased to 1 min per each kb of expected amplicons length. Each reaction set included a negative control, which lacked a template but contained all the other reaction components. After a reaction is complete, the products were resolved on 0.9% agarose gels containing 10 ug/ml ethidium bromide, visualized under UV light and the images were captured using Kodak gel imaging system (Kodak, imaging systems).

Restriction enzyme digestions

All restriction enzyme digestions were performed in a total volume of 20 µl. The reaction mix included respective 1x restriction enzyme-specific reaction buffer, 1 µg of DNA, 1-2 units of enzyme, 0.5 µl of BSA (10 µg/µl), and water to a final volume of 20 µl. Typically the restriction enzyme digestions were carried out at 37°C for 2 hours unless an enzyme requires a specific temperature. For the reactions that utilized two restriction

enzymes, a buffer optimal for both the enzymes according to the manufacturer's instructions was utilized.

Phenol purification of DNA

DNA fragments from all PCR reactions or restriction enzyme digestions were purified by phenol purification method. Typically, 3M sodium acetate was added to the final concentration of 0.3 M to a micro centrifuge tube containing DNA solution and final volume was adjusted to 200 µl with TE buffer. Two hundred micro liters of phenol (pH, 8.0) was added, vortexed and centrifuged at 12,000 g for 15 min at 4°C. The top aqueous layer was transferred into a clean micro centrifuge tube and to which 200 µl of phenol:chloroform:isoamylalcohol (25:24:1 ratio) mixture was added. The contents were mixed by vortexing and then centrifuged at 4°C at15, 000 g for 15 min. The upper aqueous layer was transferred into another clean micro centrifuge tube. These steps were repeated with the above stated pheol:chloroform:isoamylalacohol mixture and then with chloroform:isoamylalacohol (24:1 ratio) mixture. To the final recovered aqueous layer, 0.5 ml absolute cold ethanol is added, incubated at -20°C for 15 min and then centrifuged at 15,000g for 15 min. The DNA pellet was washed with 0.5 ml of 70% ethanol. The final purified DNA pellet was air dried and resuspended in 20 µl of TE buffer and stored at -20°C until further use.

Ligation reactions

The ligation reactions included approximately 25 ng of linearized purified plasmid vector DNA, 5 to10 molar excess of insert DNA, 1x ligation buffer, 5 units of T4 DNA

ligase (Promega Corporation, Madison, WI) in a 20 µl reaction volume. The ligation reactions were carried out by incubating the contents at 15°C for 16 h. Following the ligation, 1 µl of ligation mix was used for transformation by chemical method (described below). Alternatively for further use, the DNA was purified by phenol: chloroform: isoamyl alcohol method described above.

Preparation of Luria-Bertani (LB) media and LB agarose plates

To prepare 1 lit of LB liquid medium, 15 g tryptone, 10 g of yeast extract and 10 g of sodium chloride were dissolved in 1 liter of double distilled water and pH of the solution was adjusted to 7.0 by adding drop by drop of 10N NaOH. Then the medium was autoclaved at a liquid cycle. LB agar plate's preparation included similar preparation as described above but with the addition of 15 grams of cell culture grade bacto-agar prior to autoclaving. After autoclaving, the LB agar medium was allowed to cool to nearly 50°C and a desired concentration of appropriate antibiotic was added to the medium. Approximately 30 ml of medium each was poured into sterile agar plates. After solidification of the agar medium, the plates were wrapped with parafilm and stored at 4°C for subsequent use.

Preparation of *E. coli* cells for use in chemical transformation methods

The *E. coli* strains utilized to prepare competent cells were Top 10 cells (Invitrogen Technologies, Carlsbad, CA). To prepare chemical competent cells, an *E. coli* colony from a streaked plate was cultured in 3 ml of LB medium overnight in a 37°C incubator, shaking at 250 rpm. Subsequently, *E. coli* culture was re-inoculated into 100 ml LB

medium and grown in a 37°C incubator. After the cells were grown to 0.4 optical density (measured at 600 nm in a spectrophotometer), the cultures were harvested by centrifuging at 2,500 g for 5 min at 4°C. The cell pellet was resuspended in 10 ml of freshly prepared 10 mM Tris-HCl (pH, 7.5) and 50 mM $CaCl_2$, and incubated on ice for 30 min. The cultures were centrifuged again at 2,500g for 5 min at 4°C and the pellet was suspended in 2 ml of 10 mM Tris-HCl (pH, 7.5) and 50 mM CaCl2 and stored first in liquid nitrogen and subsequently at -80°C.

Transformation protocol

Transformation of ligated products into *E. coli* cells was performed by a chemical method ($CaCl_2$). Typically, 200 µl of chemical competent *E. coli* cells were mixed with 50 µl of 100 mM $CaCl_2$ and 49 µl of sterile water. One µl of ligation products were added to this suspension and mixed by gentle tapping of the tube. The contents were then incubated in ice for 15 min, followed by a heat shock at 42°C for 2 min. The cells were then incubated at room temperature for 10 min; subsequently 1 ml of LB medium was added and transferred to a 37°C shaker incubator set at 200 rpm for 1h. Cultures were then recovered by centrifugation at 2,500 g for 5 min at 4°C, resuspended in about 200 ul of the media (typically, resuspended in residual media remained after decanting the LB media after centrifugation) and plated on to LB agarose plates containing appropriate antibiotics and chemicals for selecting transformants containing plasmids with recombinant inserts. The culture was uniformly dispersed onto the agar plates using a glass rod spreader. Then the plates were incubated overnight in a 37°C incubator overnight.

Selection of recombinant clones

The presence of transformants on agarose plates containing the transformed cultures (described above) was assessed by comparing plates containing appropriate controls (ligation controls, no transformation controls). Transformants containing plasmids with inserts were selected, transferred to liquid LB media and inoculated with an appropriate antibiotic and were grown overnight at 37°C in a shaker incubator for use in isolating plasmid DNAs.

Plasmid DNA isolation

From overnight grown *E. coli* cultures, plasmid DNA was isolated by following boiling preparation method. To isolate plasmid DNA, 1.5 ml of overnight grown bacterial cultures were transferred to a 2 ml micro centrifuge tube and centrifuged 12,000 g for 5 min. The supernatant was aspirated out carefully and cell pellet was resuspended in 0.4 ml plasmid lysis buffer (10 mM Tris-HCl pH 8.0, 0.1 M NaCl, 1mM EDTA, and 5% v/v Triton X-100) with the help of a tooth pick. Twenty five micro liters of freshly prepared lysozyme (10 mg/ml) was added. (Lysozyme was prepared by dissolving 10 mg of lysozyme powder in 1 ml of 10 mM Tris-HCl buffer (pH, 8.0) to get a final concentration of 10 mg/ml.) The contents of the tube were vertexed to mix, placed in a boiling water bath exactly for 40 sec, and centrifuged at 12,000 g for 15 min at 4°C. The pellet containing cell debris was removed with the help of a tooth pick. Four hundred and twenty µl of 100%, cold (-20°C), isopropanol was added to the supernatant and mixed by vertexing, incubated at room temperature for 5 min and centrifuged at 12,000 g for 15 min to recover plasmid DNA. Supernatant was discarded and the DNA pellet was

washed with 70% ethanol and dried using a speed-vac system (Labconco Centrivap Concentrator, Kansas City, MO) typically for about 5 min. Final DNA pellets were resuspended in 100 µl of TE buffer and contaminating bacterial RNA was digested by treating with 1 µl of RNase A (1 mg/ml) at 37°C for 5 min. The presence and quality of the plasmid DNA was checked by agarose gel electrophoresis.

Isolation of genomic DNA

Genomic DNA was isolated from wild-type or from the transformed *E. chaffeensis* organisms. Genomic DNA of *E. chaffeensis* grown in macrophage cultures was isolated by sodium dodecyl sulfate-proteinase K-phenol, chloroform-isoamyl-alcohol method (336). Briefly, 1.5 ml of *E. chaffeensis* cultures were harvested by centrifugation at 12,000 g for 15 min and the cell pellet was resuspended in 0.5 ml of DNA extraction buffer (10 Mm Tris-HCl pH 8.0, 0.1 M EDTA, and 0.5% SDS) (Vijay, I think that the DNA extraction buffer also contains NaCl; please verify) containing 0.5 mg/ml proteinase K (Sigma Chemical Co., St. Louis, MO). The contents were mixed by vertexing and incubated for 6 h at 56°C. Phenol:chloroform:isoamyl alcohol extraction method (as described above) was used to remove proteinecious material and the final purified DNA was recovered by ethanol precipitation methods as described previously. DNA pellets were air dried and resuspended in 100 of TE buffer. To remove the contaminating RNA, the DNAs were treated with 1 µl RNase A (10 mg/ml) at room temperature for 10 min. The DNAs were stored at -20°C until use.

Agarose gel electrophoresis

Plasmid DNA, restriction digestion products or PCR products were analyzed by resolving them on a 0.9% agarose gels by subjecting to electrophoresis. The agarose gels were prepared by dissolving agarose powder in 1x TAE buffer (40 mM Tris-acetate, 1 mM EDTA; final pH 8.0) containing 0.1 µg/ml of ethidium bromide. The contents were poured on a gel holding device and were allowed to solidify at room temperature. The gel was placed in an electrophoresis chamber containing 1X TAE buffer with 0.1 µg/ml of ethidium bromide. About 5 µl of DNA containing 1 x gel loading buffer was loaded into the wells. Molecular weights markers were also loaded in a separate well and resolved to help in determining the approximate molecular weight of the unknown DNA fragments resolved. The DNA was subjected to electrophoresis in agarose gels at 70 V for 60-90 minutes and was visualized under UV illumination. The images are captured using Kodak gel imaging system.

Preparation of *E. chaffeensis* for transformation experiments

In vitro cultivation of E. chaffeensis

The Arkansas isolate of *E. chaffeensis* was cultivated in vitro in the canine macrophage cell line DH82 at 37°C, by following the protocols established before [87].

Purification of E. chaffeensis

E. chaffeensis infected macrophage cultures were harvested when the infectivity reached to 80-90 % in nearly 100% confluent flasks. Infection was assessed by microscopic examination of polychromatic stained cytospin slides. About 25 ml of infected culture was collected into a 50 ml sterile falcon tube and centrifuged at 15,000 g for 15 min. The supernatant was discarded and the pellet is resuspended in 10 ml of 0.25 M ice cold sucrose solution. The cells were lysed by adding grit and vortexing

twice for 30 sec. The lysed cell suspension was transferred to a new sterile tube free of glass beads and centrifuged at 100 g for 10 min. The supernatant was carefully collected and transferred to a new tube and centrifuged at 15,000 g for 15 min to collect cell-free Ehrlichia organisms. At the end of centrifugation, supernatant was discarded and the cell-free bacteria pellet was resuspended in 10 ml of 0.25 M ice cold sucrose. The centrifugation step was repeated one more time and the final *E. chaffeensis* organism pellet was resuspended in 0.6-1 ml of 0.25 M sucrose solution. Also in some experiments cell free cells were made by passing the culture through bent needle.

Preparation of plasmid DNA for electroporation

E. coli cultures containing the transposon plasmids were grown in 120 ml LB media with respective antibiotics) for about 16 h. The cultures were used to isolate plasmid DNA by following the protocols outlined by EndoFree Plasmid Purification Kit (QIANGEN, CA). Briefly, the culture was centrifuged at 6,000 g for 15 min at 4°C, and then the supernatant was discarded. The pellet was resuspended in 10 ml buffer P1 (QIANGEN, CA), and mixed thoroughly with 10 ml buffer P2 (QIANGEN, CA), followed by incubating at room temperature 24°C for 5 min. During the incubation, the QIAfilter Cartridge was prepared (screw the cap onto the outlet nozzle of the QIAfilter Maxi Catridge and place the QIAfilter Cartridge in a convenient tube). Ten ml chilled buffer P3 (QIANGEN, CA) was added to the lysate, and mixed immediately and thoroughly by vigorously inverting for 6 times (do not incubate the lysate on ice). The lysate was poured into the barrel of the QIAfilter Cartridge, and was incubated at room temperature for 10 min. The cap from the QIAfilter Catridge outlet nozzle was removed. The plunger was gently inserted

into the QIAfilter Maxi Cartridge and the cell lysate was filtered into a 50 ml tube. Two and half ml buffer ER (QIANGEN, CA) was added to the filtered lysate, mixed by inverting the tube approximately 10 times and incubated on ice for 30 min. A QIAGEN-tip 500 was equilibrated by applying 10 ml buffer QBT (QIANGEN, CA), and allowing the column empty by gravity flow. The filtered lysate was applied to the QIAGEN-tip. The QIAGEN-tip was washed with 30 ml buffer QC (QIANGEN, CA) for 2 times. Plasmid DNA was eluted with 15 ml buffer QN (QIANGEN, CA) and precipitated by adding 10.5 ml room-temperature isopropanol, followed by mixing and centrifuging immediately at 15,000 g for 30 min at 4℃. The supernatant was carefully decanted. DNA pellet was washed with 5 ml of endotoxin-free 70% ethanol stored at room temperature (40 ml of 96-100% ethanol was added to the endotoxinfree water supplied with the kit) and centrifuged at 15,000 g for 10 min. The supernatant was carefully decanted without disturbing the pellet. The pellet was air-dried for 10 minutes and the DNA was dissolved in 100 µl of endotoxin-free TE buffer (QIANGEN, CA). The purified plasmid was stored at -80℃ for later use.

Transformation of *E. chaffeensis* in DH 82 macrophage cells

E. chaffeensis cultures were grown in macrophage cell line, DH 82, and when infection reached to > 90%, cultures were harvested by spinning at 15, 000 g for 15 min at 4°C. To the pellet, 10 ml of ice-cold 0.25 M sucrose solution was added and vortexed to resuspend the culture. The host cells were broken by passing through 27.5 gauge bended needles for 4 times, and centrifuged at 200 g for 10 min at 4°C to pellet the cell debris. The supernatant containing Ehrlichia organisms was filtered through 2.7 µm filter

and spun at 15,000 g for 15 min at 4°C to collect the cell free bacteria. The pellet was washed twice with ice cold sucrose solution (0.25 M) 10 ml and the final pellet was resuspended in 0.6-1 ml of 0.25 M ice-cold sucrose. The final sucrose solution varied depending on the estimated number of organisms and 100 µl each of the purified *E. chaffeensis* suspension was used for transformation expressions. One µl plasmid DNA and purified Ehrlichia organisms were mixed in an electroporating cuvette for use in electroporation. The electroporation was performed twice at 1800 V. To serve as a control, 100 µl of cell-free bacteria were also electroporated by giving double shock at 1800 V. The electroporated cells were transferred into T 25 cm flask of DH82 cells having 80% confluency (5 ml culture volume). The culture flasks were incubated at 37°C for overnight. Fraction of cultures from each flask (1.5 ml each) were harvested at different times post transformation (6 h, 12 h, 24 h, 48 h and 6 days and till 2months in some experiments) for use in DNA isolation and analysis. To the remaining culture, media is replaced every three or four days and *E. chaffeensis* infection was monitored for up to two months. The culture flasks were maintained with an appropriate antibiotic, starting from the day 3 following transformation.

Genomic DNA isolation

The genomic DNA was isolated at many time points using the phenol and proteinase k method described in above methods.

Methods for monitoring transformants

The culture derived DNA was used to perform PCR and Southern blot analysis with specific probes.

Southern Blotting

The PCR products were resolved on a 0.9% agarose gel by subjecting to 40 V for 6 h and the resolved DNA was transferred to a nylon membrane. The prehybridization solution was prepared which contained 6 X SSC, 10 mm sodium phosphate buffer pH 6.8, 1 mm EDTA pH 8.0, 10 X Denhard's, 100 µg/ml sonicated and denatured salmon sperm DNA, and 0.5% SDS. The membrane was prehybridized for 2 h at 68℃. The ^{32}P-labeled DNA probes were synthesized by using Primer-It II Random primer labeling kit (Stratagene, Carlsbad, CA). After 14 h of hybridization at 68℃, the membrane was washed with 50 ml wash buffers; once with 6 X SSC containing 0.1% SDS followed by one wash with 2 X SSC having 0.1 % SDS, one wash with 1 X SSC containing 0.1% SDS, and a final wash was performed with 0.1 X SSC and 0.1% SDS. The nylon membrane was then exposed to an X-ray film at -70℃ by using an intensifying screen.

Semi random, two step PCR (ST-PCR)

The ST-PCR method was described by Chun et al [85]. In the ST-PCR one primer contained sequence complimentary to the transposon insertion sequence and the second primer contained a 20-nucleotide specific sequence followed by a ten bases of degenerate sequence and a specific five nucleotide sequence GATAT. A subset of these primers anneal to an unknown DNA sequence near the transposable element. These primers will allow the initial amplification. The second pair of primers is designed to be specific for the amplicons produced in the first round. One primer is complementary to the known transposon sequence and the other primer is specific to

the 20-nucleotide sequence in the semi random primer. The PCR strategy was illustrated in Figure 3.6.

Genomic DNA isolated from host cell free *E. chaffeensis* mutants was used in all ST-PCRs as templates. In particular, the cultured organisms were purified by following renografin purification method to eliminate host cells and contaminating host genomic DNA. Briefly, the mutant *E. chaffeensis* infected DH82 cultures were harvested when the infection level reached to ~ 90%. The cells were harvested and pelleted by spinning at 12,000 g. The pellet was suspended in 1x PBS and was passed 4-5 times through a bent needle to break the DH82 cells for releasing *E. chaffeensis* organisms. The suspension was then filtered through 2.7 um and 1.6 um filters to recover host cell free *E. chaffeensis* organisms were suspended in 1 x PBS. For the renografin gradient, different concentrations of renografin (34%, 44% 54%) were prepared by diluting the stock renografin in 1x PBS. Gradients were made in a clear ultracentrifuge tubes, the *E. chaffeensis* organisms in PBS were layered above the 34% renografin layer. The gradient tubes containing the organisms were centrifuged at 30,000 rpm for 1 h *Ehrlichia* organisms remain at the interfaces between the PBS and 34% renografin; 34% and 44% renografin; 44% and 54% renografin due to differences in the densities of the organisms. Bacterial fractions from all three interfaces were collected, pooled and diluted in PBS and centrifuged at 25,000 rpm for 30 min to recover the organisms. The resulted *Ehrlichia* pellet was used to isolate genomic DNA by phenol/chloroform/isoamyl alcohol extraction method for use as the template for the ST-PCR reactions.

In the first round of ST-PCRs, 10 ng each of host cell-free *Ehrlichia* genomic DNA was used in 20 ul reaction mixture containing 50 mM KCl, 10 mM Tris HCl (pH 8.3), 200

ug/ml gelatin, 3 mM $MgCl_2$, 200 uM of dNTPs, 20 pmol each of PCR primers. (Primer set 1 and 2 for first reaction; 3 and 4 for 2nd reaction.) ST-PCRs were performed using BIORAD Real time PCR system. The products from the first PCR were resolved on an agarose gel and subsequently the products were diluted 1:100 in TE buffer and one micro liter each was used as the template for the second round of PCR by following the conditions optimized for the assays.

Table 1 ST-PCR primers

SET	PRIMERS
1	RRG 1260 (random for 1st round)
2	RRG 1194, RRG 1259, RRG 1264
3	RRG 1261(random for 2nd round)
4	RRG 1258, RRG1194, RRG1202

Table 2*ST –PCR first round settings*

1	**94^0C , 2min**
2	94^0C, 30 seconds
3	Initial tem = 42^0C, 30 seconds -1.^{0}C for each subsequent cycle
4	72^0C, 3 minutes
5	GOTO 2 5x more
6	94^0C, 30 seconds 65^0C, 30 seconds 72^0C, 3 minutes
7	GOTO 6 24x more
8	4^0C , end

Table 3: ST-PCR second round settings

1 940C, 30 seconds

2 650C, 30 seconds

3 720C, 3minutes

GOTO 1, 29 x more

4 40C

5 END

DNA sequencing

The products from the ST-PCR reactions were sequenced on ceq sequencer. The protocols for sequencing analysis were followed as described in the CEQ 2000 Dye Terminator Cycle Sequencing with Quick Start Kit manual (Beckman Coulter, CA).

BLAST search analysis

The sequences obtained from the ST-PCR products were searched against the *E. chaffeensis* genome available at NCBI database by performing BLAST search program available at the NCBI website.

Chapter 3 - Results

Transposon mutagenesis of *E. chaffeensis*

Recent studies have shown that Himar 1 transposase system is useful in creating insertion mutations in both Gram positive and Gram negative bacteria [81, 83] This method is also used in creating mutations in intraphagosomal, pathogenic bacteria such as *A. phagocytophilum, A. marginale* and *C. burnetti* [79, 83] Our research team collaborated with Dr. Ulrike Munderloh at University of Minnesota in creating random mutations in *E. chaffeensis* by utilizing transposon- mutagenesis method [88]. The Himar 1 transposon plasmid constructs previously used for mutagenesis of *A. phagocytophilum* were modified to create a single plasmid system. The plasmid construct contained both transposase coding sequence and transposon insertion segment with transposase-specific recognizable inverted repeats are engineered into single plasmid. The transposable element contained either the GFPuv or mCherry expression cassette, as well as the spectinomycin resistance gene cassette (*aad* gene) and their expression is driven by *A. marginale* transcription regulator promoter (*Amtr*). Similarly, the transposase expression is driven by the *Amtr* promoter. Three independent experiments were performed; one with the mCherry plasmid and two with GFPuv plasmid (plasmids are illustrated in Figures 3.1 and 3.2). The transformation experiments were performed in *E. chaffeensis* organisms recovered from ISE6 tick cell cultures as described earlier [81]. The transformants were propagated in initially in ISE6-tick cells to select mutants conferring resistance to spectinomycin and streptomycin (100ug/ml each) and then transferred to DH82 cells while maintaining the antibiotics in the culture media. The cultures were also observed microscopically for the

expression of mCherry or GFPuv. My research involved mapping the genomic sites from the mCherry mutants and the GFP mutants from one experiment each.

Southern blot analysis

The transposon insertions in the mutants were verified by performing Southern blot analysis (Figures 3.3, 3.4 and 3.5). Genomic DNA isolated from the transformed cultures of *E. chaffeensis* was digested with restriction enzymes which do not have recognition sites to generate single DNA fragments recognizable when probed with the probe used for DNA blot hybridization. About 100 ng each of DNA isolated from the transformed cultures was digested with the following restriction enzyme; Hind III, Bgl II, Spe I/Nde I, Bsrg I or EcoR V/Bgl II. Digested DNAs were resolved on an agarose gel, transferred to a nylon membrane, and probed with the ^{32}p-labeled *aad* gene segment. Multiple restriction digested DNA segments were recognized by the probe in the DNAs of mutants derived from the transposon constructs containing mCherry and GFPuv (Figure 3.1 and 3.2).

Southern blot analysis was also performed using the mutants' genomic DNA isolated at two different time points following their growth to assess the stability of the mutants in culture. The Sothern blot data revealed variations in the DNA fragments recognized by the *aad* probe suggesting that the genomic DNA from the cultured mutants represented pool of multiple mutants and that the mutants' growth is variable during their growth in culture.

Mapping of the transposon insertion sites using ST-PCR

For mapping the exact location of the transposon insertions in the *E. chaffeensis* chromosome, we used the ST-PCR method described by [85] (Figure 3.7 has a cartoon showing the ST-PCR strategy). Three independent ST-PCR experiments were performed using different primer combinations (Figures 3.7-3.13). The final products from the 2nd round ST-PCR were resolved on an agarose gel and the entire strong and discrete DNA bands were recovered from the gels and the DNA were purified. The DNAs were subjected to sequence analysis. The resulting sequences were compared with the *E. chaffeensis* genome data to identify the location of the insertion sites. The alignment data identifying the insertion sites are shown in Figure 3.14.

The analysis aided in the identification of four transposon insertion locations in the genome. Further, it verified four other insertion sites previously identified by rescue cloning method. Together, the ST-PCR method was valuable in confirming the location of transposon insertions at 8 genomic sites within the *E. chaffeensis* chromosome. Six of the 8 transposon insertions were present within the non-coding regions of the genome and the remaining two insertions were present within the coding regions of two hypothetical proteins (Ech_0379 and Ech_0601) (Figure 3.14).

Transposon insertion regions

MCherry mutant

1. Inside non coding region between the genes ECH_0230 and ECH_0231
2. Inside non coding region between the genes ECH_0284 and ECH_0285

3. Inside non coding region between the genes ECH_0479 and ECH_0480
4. Inside non coding region between the genes ECH_0490 and ECH_0492
5. Within the coding sequence of the gene ECH_0379

GFPuv mutant

1. Inside non coding region between the genes ECH_0202 and ECH_0203
2. Inside non coding region between the genes ECH_0760 and ECH_0761
3. Within the coding sequence of the gene ECH_0601

Some of the BLAST program alignments performed in mapping the insertion mutations

Mapping Insertion at 478108 (in between Ech_0490 (forward) and Ech_0492 (complement)) as 478107-ITR-AmTr-mCherry-Spec-ITR-478108

Sequence given below is derived from ST-PCR second round amplicon

TGCTTTACGCGAGGGTGGGCGGATATCAGACTTATAAAACGTATTAAGAAATCTTAC
AAGCATTTATAATAAAATAATATATTAGTCGATACCATTGTAAAAATGGCTCAACGT
TTACCTTTTTACCTCAACCTCACATAATATATAATCAGATTTATATCTAGATTCTATT
TTTCTATTAAATATACAAAAAATTTACTTCTAACAGGTTGGCTGATAAGTCCCCGGT
CTAAGCTTGCATGCCTGCAGGTCGACTCTAGATTATTTGCCGACTACCTTGGTGATCT
CGCCTTTCACGTAGTGAACAAATTCTTCCAACATGA

The sequence was blasted against the E. chaffeensis str Arkasnas genome using the BLAST algorithm. The alignment shows that part of the sequence is complimentary to the E. Chaffeensis genome and other part is complimentary to the transposon insertion sequence. This gives the data regarding the point of insertion and the exact location of the insertion within the genome

Alignment

> gb|CP000236.1| Ehrlichia chaffeensis str. Arkansas, complete genome
Length=1176248

Sort alignments for this subject sequence by:
E value Score Percent identity
Query start position Subject start position

Features flanking this part of subject sequence:
165 bp at 5' side: lipoic acid synthetase (Ech_0490: 477049-477942)
253 bp at 3' side: putative phosphate ABC transporter, permease protein (Ech_0492:478543::479811)

Score = 327 bits (362), Expect = 8e-91
Identities = 183/184 (99%), Gaps = 0/184 (0%)
Strand=Plus/Minus

```
Query  22      GATATCAGACTTATAAAACGTATTAAGAAATCTTACAAGCATTTATAATAAAATAATATA  81
               ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  478290  GATATCAAACTTATAAAACGTATTAAGAAATCTTACAAGCATTTATAATAAAATAATATA  478231

Query  82      TTAGTCGATACCATTGTAAAAATGGCTCAACGTTTACCTTTTTACCTCAACCTCACATAA  141
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  478230  TTAGTCGATACCATTGTAAAAATGGCTCAACGTTTACCTTTTTACCTCAACCTCACATAA  478171

Query  142     TATATAATCAGATTTATATCTAGATTCTATTTTTCTATTAAATATACAAAAAATTTACTT  201
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  478170  TATATAATCAGATTTATATCTAGATTCTATTTTTCTATTAAATATACAAAAAATTTACTT  478111

Query  202     CTAA  205
               ||||
Sbjct  478110  CTAA  478107
```

The above sequence was matched with the E. chaffeensis genome

Sequence matched with transposon sequence (ITR-AmTr-cherry-Spec-ITR seq)

```
Score =  212 bits (234),  Expect = 7e-59
 Identities = 117/117 (100%), Gaps = 0/117 (0%)
 Strand=Plus/Minus

Query  203   TAACAGGTTGGCTGATAAGTCCCCGGTCTAAGCTTGCATGCCTGCAGGTCGACTCTAGAT  262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1843  TAACAGGTTGGCTGATAAGTCCCCGGTCTAAGCTTGCATGCCTGCAGGTCGACTCTAGAT  1784

Query  263   TATTTGCCGACTACCTTGGTGATCTCGCCTTTCACGTAGTGAACAAATTCTTCCAAC  319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1783  TATTTGCCGACTACCTTGGTGATCTCGCCTTTCACGTAGTGAACAAATTCTTCCAAC  1727


 Score = 53.6 bits (58),  Expect = 4e-11
 Identities = 29/29 (100%), Gaps = 0/29 (0%)
 Strand=Plus/Plus

Query  203  TAACAGGTTGGCTGATAAGTCCCCGGTCT  231
            |||||||||||||||||||||||||||||
Sbjct  1    TAACAGGTTGGCTGATAAGTCCCCGGTCT  29
```

```
LOCUS       CP000236                2763 bp    DNA     linear   BCT 11-MAR-
2010
DEFINITION  Ehrlichia chaffeensis str. Arkansas, complete genome.
ACCESSION   CP000236 REGION: 477049..479811
FEATURES             Location/Qualifiers
     source          1..2763
```

The italicized sequences are both upstream and downstream of the open reading frame. The arrow shows the exact position of the transposon sequence insertion.

```
   1 atgagaagta aacctgattg gttaaaagtt aagatgccta caggggatac tttttatcaa
  61 gttcgcaatt taatgaaatt gtataaatta aatactgtgt gtgaagaagc agcttgtccc
 121 aacattggtg agtgttggaa taaaagacat gcaacagtaa tgattttagg ttctacatgt
 181 actagagcat gtgcattttg taatgttgtt tctggaattc cagataaact ggatcctcat
 241 gagcctcaaa acttagctaa ggctgttggt ttacttaaac ttgagcacgt tgttattaca
 301 tctgtagaca gggatgactt ggaagatggt ggatctggac attttgtaga gtgtatagaa
 361 gaaatacgaa aaaacgatca gaatgttact attgaagtat taactccaga ttttttgaat
 421 aaacatggtg caattgaaaa ggttgctgat gctgctcctg atgtttacaa tcataatata
 481 gaaactgttc caaggttgta tgcaaaaatt agaccaaaag cacgttactt tcattctctt
 541 tatttgttaa aaacagttaa atataaaaat cctaaagtat ttactaaatc tggaattatg
 601 gtaggattgg gtgaaacaaa agaagaaata tatcaagtaa tgaatgattt aagatctgcg
 661 gatgttgatt ttataacgat tggtcaatat ttgcagccta ctcctaagca tgctgcagtt
 721 gataggtatg taactcctga agaatttgat cattacaagt atgttgcata ctctaaaggg
 781 tttttaatgg ttgcgtcggg ccctttagta aggtcttcat accatgctgg agaggatttt
 841 caaaggttaa agaaaaatcg tgctgctatg ttcatgcatg ctaaaagcaa ttagaattaa
 901 catggtagtt aatatttttt agaagtttga atgtttatac cagaagggtt tttcattagt
 961 tttatatgat atttaatttt ttagatagtt aatcggtagt gtagttggtt gtaatctgag

                                                Site of Insertion
                                                        ⇩
1021 gaaagttttt gtgattgaga gaaagtaggc tatacata ttagaagtaaat tttttgtata
1081 tttaatagaa aaatagaatc tagatataaa tctgattata tattatgtga ggttgaggta
1141 aaaaggtaaa cgttgagcca tttttacaat ggtatcgact aatatattat tttattataa
1201 atgcttgtaa gatttcttaa tacgttttat aagtttgata tcgcacttat agctcagtgg
1261 atagagcgtt accctccgga ggtaaaagtc gcaggttcaa atcctgctaa gtgcacttat
1321 tttttatgtt attataaaat attaatgtgg agcaacaata ttgttgatgt tatctctttt
1381 ttcatagttt ggtaatgata gtttagttaa tgttgcttta tgattattaa tatatatggg
1441 gaattattat ttgtaatgta ttaattgaag agtgaatcat tgaatttaag tattttaaaa
1501 aataaaatga tcaaactttt tccttataaa agttactata agatttagaa ttaataacat
1561 taaaagtaga gctactatag caattgcagc taattggata aattcaatat gtggattact
1621 tgcccacata tatatttgta caggtaatac agttgttgga tcgagaaaat cctttggaat
1681 atcaactatg aatgctacca tcccaatcat tattaacgga gaagattcac ctaagattct
1741 tgcaatactt aaaatagcac cttgcatgat actaggtaat gctattggaa atgagtgatg
1801 tagaattacc tgcatatttg atgctcctag tgagaaagct gcatctttta cggatggtgg
1861 cactgctgca aatgcttgtt ttgtagatgt taccaaattt ggtaacatca ttaatgatag
1921 agtcatacca cctataagtg ctgatgaacg tggtatatca aaaatattga tgtatattgc
1981 taaccctaaa atgccaaata ttattgacgg tacagatgca agatttgtta tgctgatttc
2041 gataatagaa gatattttat tttttggtat gagctcactt aaacatatgc cggatgtgat
2101 acctataggt aatgcgaatg ttaaacatat aattattgta aacatggatc ctattagtga
2161 gccaagaatc ccagtatttt caggagatct agagttatat tttgtaaaca gtgcgtcatt
2221 aaatgatact tttatcctgt tttcattttg taattttgca ataacttttt gatatattgc
```

```
2281 atgatttatg tttttgtttt tgatagcttt attaaatgta gtggaagatt gaaaccatat
2341 gttgaattta gtggattttg gattcttaag tagaaaatct gctaaattat gatgagcagt
2401 atcatttatt atattataaa tgatttttct attgtcagga tttgaaagat ctataacatt
2461 tgaaaatagt tcctctaagg attgatgaat aaaatatatt agattgtttc tcaatgttag
2521 tttatcttgt tgttctatag tatcattatt gattgtaatt gataataaaa tttttgttgt
2581 tgttagtgca ttatatccat tgttaacgat acttcctaaa attataaaaa gtatacctaa
2641 ggaactaaca agtgcagtca ttgataggat tgataagata cgatccttcc tatttttaa
2701 ggatatacgg gaagcaatat gacgatactt tagtaattta cctaattttt ttcgtatgtg
2761 cat
```

Mapping of Ech_0202

Sequence obtained from ST-PCR

GCGAGAGCGTGAGGCGGTATATAAAGATTTAAAGCGAATACAGTAATGGTATTAGT
TTTTCGTAAAATATGAGAGGAGCATGTTCAGGGTGTCCAAGTGCCTCAATTACTCTG
AAAAGAAGGTATATATAATATGTTGAGCTATTATATTCCGGATATACAAGGTGTAGA
GAGTATTCAGTAGAAATAGAATATATTAGTTTTGCTGTTGTGTTGGAAAGTACTTTTT
AAGTTTTTAAGTCTATAGAAATTATCTGTAACAGGTTGGCTGATAAGTCCCCGGTCT
AAGCTTGCATGCCTGCAGGTCGACTCTAGATTATTTGCCGACTACCTTGGTGATCTC
GCCTTTCACGTAGTGAACAAATTCTTCTCTTTTTCATGAATTAAAACAAATACTTTTT
TTTTAGA

Blasted against *E. chaffeensis* genome

```
>    gb|CP000236.1|  Ehrlichia chaffeensis str. Arkansas, complete genome
Length=1176248


 Score =  363 bits (402),  Expect = 1e-101
 Identities = 230/240 (96%), Gaps = 8/240 (3%)
 Strand=Plus/Plus

Query  19      ATATAAAGATTTAAAGCGAATACAGTAATGGTATTAGTTTTTCGTAAAATATGAGAGGAG  78
               ||||||||  |||||||   |||||||||||||||| |||||| |||||||  |||||||||
Sbjct  193038  ATATAAAG-TTTAAAGG--ATACAGTAATGGTATT-GTTTTT-GTAAAAT-TGAGAGGAG  193091

Query  79      CATGTTCAGGGTGTCCAAGTGCCTCAATTACTCTGAAAAGAAGGTATATATAATATGTTG  138
               |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Sbjct  193092  CATGTTCAGGGTGTCCAAGTGCCTCAATTACTCTGAAA-GAAGGTATATATAATATGTTG  193150

Query  139     AGCTATTATATTCCGGATATACAAGGTGTAGAGAGTATTCAGTAGAAATAGAATATATTA  198
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  193151  AGCTATTATATTCCGGATATACAAGGTGTAGAGAGTATTCAGTAGAAATAGAATATATTA  193210

Query  199     GTTTTGCTGTTGTGTTGGAAAGTACTTTTTAAGTTTTTAAGTCTATAGAAATTATCTGTA  258
               |||||||||||||||||||||||||||||||||||||||||||||| | ||||||||||
Sbjct  193211  GTTTTGCTGTTGTGTTGGAAAGTACTTTTTAAGTTTTTAAGTCTATA-AGATTATCTGTA  193269
```

193269-ITR-Spec-GFP-AmTr-ITR-193270

Sequence matched with transposon sequence (to ITR-AmTr-cherry-Spec-ITR seq)

```
Query  257   TAACAGGTTGGCTGATAAGTCCCCGGTCTAAGCTTGCATGCCTGCAGGTCGACTCTAGAT  316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1843  TAACAGGTTGGCTGATAAGTCCCCGGTCTAAGCTTGCATGCCTGCAGGTCGACTCTAGAT  1784
```

```
Query  317   TATTTGCCGACTACCTTGGTGATCTCGCCTTTCACGTAGTGAACAAATTCTTC  369
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1783  TATTTGCCGACTACCTTGGTGATCTCGCCTTTCACGTAGTGAACAAATTCTTC  1731
```

```
 Score = 53.6 bits (58),  Expect = 5e-11
 Identities = 29/29 (100%), Gaps = 0/29 (0%)
 Strand=Plus/Plus

Query  257  TAACAGGTTGGCTGATAAGTCCCCGGTCT  285
            |||||||||||||||||||||||||||||
Sbjct  1    TAACAGGTTGGCTGATAAGTCCCCGGTCT  29
```

Mapping of Ech-0480

Sequence obtained from ST-PCR:

ATATATATAATAATATGTTATAATTTAATTATATTGTATAATCATATGATAATATAGC
TTTTTGTTCTAATTTATATTCTAAAGTAACAGGTTGGCTGATAAGTCCCCGGTCTGAA
TTCGGCTCCTCCTAGAACGATCGCCGCATGCTAGCATAAAACGCGCGCACTAACAA
GTGCCCAGCCGCGCTGCCTCAAAAAATCTCCGGAAGTTTCCCGGATATTTATCGATG
ATTTTATTCTTTAATCTAGCATACCATTTACAGTTCCTATTATTACAACCATACATTA
AATAACTGTTGCATTAACATTATGTATGATTTATCCTAAGTTATCTGAGTAACATAA
GGGATAACCNCTAAATTTTTGTCTCNTTTCACCAACCCCCCAATTCGGGGCCGGG

Blasted against the Ehrlichia chaffeensis str. Arkansas, complete genome

```
Query  85   TAACAGGTTGGCTGATAAGTCCCCGGTCTGAATTCGGCTCCTCCTAGAACGATCGCCGCA  144
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    TAACAGGTTGGCTGATAAGTCCCCGGTCTGAATTCGGCTCCTCCTAGAACGATCGCCGCA  60

Query  145  TGCTAGCATAAAACGCGCGCACTAACAAGTGCCCAGCCGCGCTGCCTCAAAAAATCTCCG  204
            |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Sbjct  61   TGCTAGCATAAAACGCGCGCACTAACAAGTGCCCAG-CGCGCTGCCTCAAAAAATCTCCG  119

Query  205  GAAGTTTCCCGGATATTTATCGATGATTTTATTCTTTAATCTAGCATACCATTTACAGTT  264
            |||||||||||||||||| |||||||||||| ||||||||||||||||||||||||||
Sbjct  120  GAAGTTTCCCGGATATTTATTGATGATTTTATT-TTTAATCTAGCATACCATTTACAGTT  178

Query  265  CCTATTATTACAACCATACATTAAATAACTGTTGCATTAACATTATGTATGATTTATCCT  324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  179  -TTATTATTACAACCATACATTAAATAACTGTTGCATTAACATTATGTATGATTTATCCT  237

Query  325  AAGTTATCTGAGTAACATAAGGGATA  350
            |||||||||||||||||||||||||
Sbjct  238  AAGTTATCTGAGTAACATAAGGGATA  263
```

Mapping of Ech 0760

Sequence obtained form the second round ST-PCR DNA band

TCGTAAATGACAATGTAATGAGATATCAGCTAGATAGATGAAACACATGTGTGTAT
AACTATGAGTAATGAATATACATGATGATGATACAATAACTATCTAGCCGTGGTGTT
TATGTAATCTAGGATGAACTAATGATGACTGCTAGTGTGTATGATGATTATATACCA
TCAGTAAAATATGTAATAAAAATCAATAGACACAGGTCATCATGATATAAAAACTG
ATAATAGAACTAAGCCTCTAACAAAAACTCATATAAAATGTTGGTAACATATCAAA
ATTAAAAAACACTTCAAATCATGAAAGCATTAATAAAGTTTCTCTCAAATTATCACT
AACTAACAGGTTGGCTGATAAGTCCCCGGTCTGAATACGGCTCCTCCTAGAACGATC
GCCGCATGCTAGCATAAAACGCGCGCACTAACAAGTGCCCAGCGCGCTGCCTCAAA
AAATCTCCGGAAGTTTCCCGGATATTTATTGATGATTTTATTTTTAATCTAGCATACC
ATTTACAGTTTTATTATTACAACCATACATTAAATAACTGTTGCATTAACATTATGTA
TGATTTATCCTAAGTTATCTGAGTACATGGAGAAAAAAAAAGGAAAAAAAATTAAG
GGGTAAATGGAAAAANGGTAAATTAAAAAGGGTAAGGGGGGGGGTAAAAAAAAAT
AAAAAAAGGGAAAAAAGGGGGGAGTGGCCCAGAAAATAAAAAA

>gb|CP000236.1| Ehrlichia chaffeensis str. Arkansas, complete genome
Length=1176248

Sort alignments for this subject sequence by:

E value Score

Percent identity

Query start position Subject start position

Features flanking this part of subject sequence:
80 bp at 5' side: RNA polymerase sigma factor RpoD
150 bp at 3' side: DNA primase

Score = 269 bits (298), Expect = 4e-73
Identities = 211/236 (89%), Gaps = 16/236 (7%)
Strand=Plus/Plus

```
Query  109     GTGTTTATGTAATCTAGGATGAACTAATGATGACTGCTAGTGTGTATGATGATTATATAC  168
               |||||||| ||| || |||| ||||||||||   |  |||||| |||   ||||| | ||
Sbjct  768139  GTGTTTAT-TAA-CT-GGAT-AACTAATGATTTTT--TAGTGT-TATA--GATTAAA-AC  768188

Query  169     CATCAGTAAAATATGTAATAAAAATCAATAGACACAGGTCATCATGATATAAAAACTGAT  228
               |||||||||||||  |||||||||||| |||||||  ||||||||||||||||||||||
Sbjct  768189  CATCAGTAAAATAT--AATAAAAATCAA-AGACACAT-TCATCATGATATAAAAACTGAT  768244

Query  229     AATAGAACTAAGCCTCTAACAAAAACTCATATAAAATGTTGGTAACATATCAAAATTAAA  288
               || |||||||| ||||||||||||||||||||  | ||||||||||||||||||||||
Sbjct  768245  AA-AGAACTAAGTCTCTAACAAAAACTCATATAACTT-TTGGTAACATATCAAAATTAAA  768302

Query  289     AAACACTTCAAATCATGAAAGCATTAATAAAGTTTCTCTCAAATTATCACTAACTA  344
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  768303  AAACACTTCAAATCATGAAAGCATTAATAAAGTTTCTCTCAAATTATCACTAACTA  768358
```

Sequence matched with transposon sequence (ITR-AmTr-Cherry-Spec-ITR)

```
Query  343  TAACAGGTTGGCTGATAAGTCCCCGGTCTGAATACGGCTCCTCCTAGAACGATCGCCGCA  402
            |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Sbjct  1    TAACAGGTTGGCTGATAAGTCCCCGGTCTGAATTCGGCTCCTCCTAGAACGATCGCCGCA  60

Query  403  TGCTAGCATAAAACGCGCGCACTAACAAGTGCCCAGCGCGCTGCCTCAAAAAATCTCCGG  462
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61   TGCTAGCATAAAACGCGCGCACTAACAAGTGCCCAGCGCGCTGCCTCAAAAAATCTCCGG  120

Query  463  AAGTTTCCCGGATATTTATTGATGATTTTATTTTTAATCTAGCATACCATTTACAGTTTT  522
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121  AAGTTTCCCGGATATTTATTGATGATTTTATTTTTAATCTAGCATACCATTTACAGTTTT  180

Query  523  ATTATTACAACCATACATTAAATAACTGTTGCATTAACATTATGTATGATTTATCCTAAG  582
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181  ATTATTACAACCATACATTAAATAACTGTTGCATTAACATTATGTATGATTTATCCTAAG  240

Query  583  TTATCTGAGTA  593
            ||||||||||||
Sbjct  241  TTATCTGAGTA  251


 Score = 53.6 bits (58),  Expect = 1e-10
 Identities = 29/29 (100%), Gaps = 0/29 (0%)
 Strand=Plus/Minus

Query  343   TAACAGGTTGGCTGATAAGTCCCCGGTCT  371
              ||||||||||||||||||||||||||||
Sbjct  1837  TAACAGGTTGGCTGATAAGTCCCCGGTCT  1809
```

Mapping of ECH-0379

Sequence obtained from ST-PCR

TGACCTTTTGAGTTATGACACGCTAGTGCATATCAAGTGTAAGCATGTAGTGGTATT
ATCTATATAATAAGTAATAGTATATTACTATCCTAATAGTAATATGGGGAGTATTAG
ATAAGCTTTGTTTTGATCTATGTATTTAGATCCTCAGGATTTTGTAACACAAGTTATA
CGATTCTAGTGCATGTTTTTAGCATCTTTATATTCGAATATACTTGGATCTTAGGTTT
AGTACAATTTCTTGATTAGAAATGATGATATCTTGTGCCACATTCAGACAACACTAA
TTGCTAAGTATAAGCTATTAAGCCTATAAGCATTAAGTGTAACATTATCGCAAGTCT
AGACTGGCATGGCATTATACCAGGGTTATGGCTTGATACAGTACCCCGGATCTAATG
CTTAGCGATGCCTGCAGAGTCGACTCTAGCATGTATGTATGCTCGACTACCTGTAGC
GTGATCTGCGCCATTGTACACGATAGTGTAACTAGACATTACTGTCCATACTTGGAT
TTCTGGCGGCCTGACGAAGGGGCCAAATGATGAATTCGTTTCATATGTTCAACAAAC
GAATAAACGCACTTGTCCTTATGCTTTCATATGGTATTCGAACTGGGACTCGACTAC
GTCGCGGCCTGGGCAATAGTCCGCNTCCCAGTTTCGCTCACATGTTCAGAGGCATAC
TTGATCCAGTGCCCTTTACTGGGCTGCTGAACTTATTTAGNCCAGTGGTCTATCTTGG
CAGCATAGATACCTGACAAGTAGTCCGAGGGATCATACTGTTAAATGCCACTTTACC
ATTTAATCGGCGTCAATTCGGCTCTAGACTCCCATGTTCGGGGGTATTGTANGAACT
TACACTATTATGCTGTTTAAGATGGGATTTTTCAGTTCTTTAGCCAGTCCCATCAAGA
ATTTAAGGAATTCCGGTGGGTTACAAATTGAATACCTCGGGAAATTCNATATAGATA
ATGNTTTCCCCTTGCAAGTTCAAGGCTTGGGAAACCTGT

```
>  gb|CP000236.1|  Ehrlichia chaffeensis str. Arkansas, complete genome
Length=1176248


  Subject start position
 Features in this part of subject sequence:
   hypothetical protein

 Score =  154 bits (170),  Expect = 3e-38
 Identities = 220/278 (79%), Gaps = 40/278 (14%)
 Strand=Plus/Plus
```

```
Query  24      TAGTGCATATCAAGTGTAAGCATGTAGTGGTATTATCTATATAATAAGTAATAGTATATT  83
               |||||||||| ||||||||||||||||||  ||   ||| ||||||| |||||   |||
Sbjct  374192  TAGTGCATAT-AAGTGTAAGCATGTAGTGGATTT---TAT-TAATAAG-AATAG---ATT  374242

Query  84      ACTATCCTAATAGTAATATGGGGAGTATTAGATAAGCTTTGTTTTGATCTATGTATTTAG  143
                || |||||  |   |||||||||| ||| ||||||||||||||||||||   ||||||
Sbjct  374243  TCT-TCCTATGA---ATATGGGGAG-ATTTGATAAGCTTTGTTTTGATCTA--GATTTAG  374295

Query  144     ATCCTCAGGATTTTGTAACACAAGTTATACGATTCTAGTGCATGTTTTTAGCATCTTTAT  203
               ||||||| |||||||  | |||||||||| |||| || |||| |||||  || || ||||
Sbjct  374296  ATCCTCA-GATTTTG--AAACAAGTTATA-GATT-TA-TGCA-GTTTT--GC-TC-TTAT  374344

Query  204     ATTCGAATATACTTGGATCTTAGGTTTAGTACAATTTCTTGATTAGAAATGATGATATCT  263
               ||| ||| ||  |||||||||| | ||||||||||||||  | ||||||| ||| |||
Sbjct  374345  ATTTGAA-AT-TTTGGATCTTAAG-TTAGTACAATTTCTTG--TTGAAATGA-GAT-TCT  374397

Query  264     TGTGCCACATTCAGACAACACTAAT-TGCTAAGTATAA  300
                 | ||||||    | ||||||||| |  | |||||||
Sbjct  374398  --TTCCACAT---CAAAACACTAATGTAATTAGTATAA  374430
```

Sequence matched with transposon sequence- ITR-AmTr-Cherry-Spec-ITR:

```
Query  373   TGGCTTGATACAGTACCCCGGATCTAATGCTTAGCGATGCCTGCAGAGTCGACTCTAGCA  432
             ||||| |||| ||| |||||| ||||| |||| || |||||||||| |||||||||| |
Sbjct  1829  TGGCT-GATA-AGT-CCCCGG-TCTAA-GCTT-GC-ATGCCTGCAG-GTCGACTCTAG-A  1779

Query  433   TGTATGTATGCTCGACTACCT  453
             |   | | ||| |||||||||
Sbjct  1778  T---TATTTGC-CGACTACCT  1762
```

Analysis was done similarly in mapping the other insertions described in figure 3.14.

Genomic insertion region verification by PCR and sequencing analysis

For verification of the insertion sites mapped through ST-PCR and sequencing, we designed primer sets for the amplification of a segment from all 8 genomic insertion sites. In each primer set, one primer was designed to bind the 5' or 3' end of the inserted transposon sequence and the second primer was designed to bind to the genomic region upstream or downstream of the insertion sites. Genomic DNA isolated from the mutants was used in amplification (Figure 3.15). The resulting PCR products were subjected to automated DNA sequence analysis. These analyses verified the results of ST-PCR. Specifically, all 8 insertion site-specific amplifications resulted in the

generation of the predicted amplicons which also contained the DNA sequence as identified by the ST-PCT and sequence analysis.

Modifying the transposon plasmids for improving the efficiency of creating transposon mutations in *E. chaffeensis*

The plasmids pHIMAR1-UV-SS (transposon plasmid with Amtr promoter GFPuv and spectinomycin resistance gene) and pET28AMTR-A7-HIMAR (transposase expression plasmid) represent a two-plasmid set used for creating transposon mutagenesis in the rickettsial pathogen, *A. phagocytophilum* [81]. This plasmid pair is similar to the single plasmid containing both the transposon insertion sequences and transposase gene sequences which were used in creating mutations in *E. chaffeensis*. Both the systems, however, use the *Amtr* promoter and spectinomycin resistance gene, *aad*. Although we succeeded in creating mutations with the single plasmid system containing the *Amtr* and *aad* gene, we reasoned that the use of a constitutively expressed *E. chaffeensis* promoter may be more efficient in creating mutations with higher efficiency. Secondly, as spontaneous mutations are possible which confer resistance against spectinomycin, we reasoned that it is desirable to use an alternate antibiotic resistance cassette such as the chloramphenicol acetyl transferase resistance gene (CAT gene) in place of *aad* gene. Therefore, I focused to optimize the two plasmid system by replacing the *Amtr* promoter with the *E. chaffeensis* transcription elongation factor gene promoter (*Ech-tuf* promoter) and *aad* gene with CAT gene in the pHIMAR1-UV-SS plasmid. The *Ech-tuf* promoter was selected because it is constitutively expressed in *E. chaffeensis* during its growth in both vertebrate and tick cell environment (unpublished results of Cheng et al.). Moreover, we reasoned that it

serves as an efficient promoter because this is involved in transcribing several genes involved in the protein biosynthesis. Similarly, we opted CAT gene in place of *aad* gene, because spontaneous mutations are less likely arise against chloramphenicol. Standard molecular manipulations of restriction enzyme digestions, cloning and sequence evaluations were performed on the pHIMAR1-UV-SS plasmid to replace the *Amtr* promoter and the aad gene (Figures 3.16-3.21). The modified plasmid was named as pHimar UV-SS cat 224 .This plasmid is now ready for use in transposon experiments to be performed on *E. chaffeensis*.

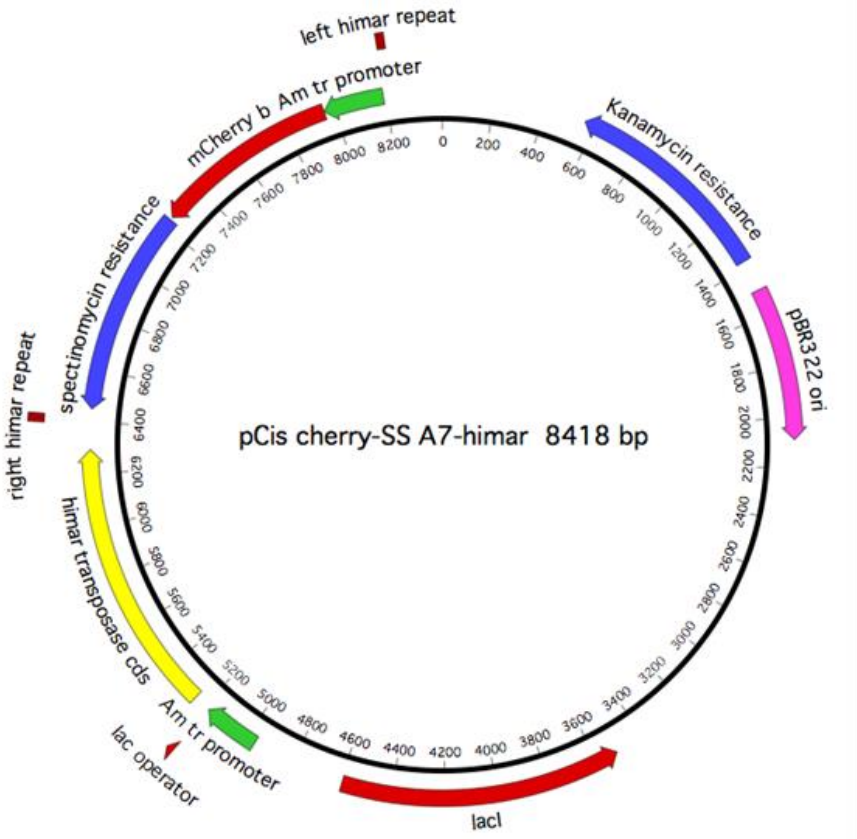


Figure 3.1 pCis mCherry –SS A7 –Himar plasmid map (prepared by Munderloh's group) pCis cherry –SS A7 –himar plasmid containing the Himar 1 transposase gene and a transposon segment having the gene sequences encoding for mCherry and the spectinomycin resistance gene (*aad* gene). The transposon segment is flanked by the left and right himar inverted repeats to facilitate recognition of the transposase enzyme. To facilitate transposase, mCherry and add gene expression, the A. marginale AmTR promoter is inserted in front of these genes.

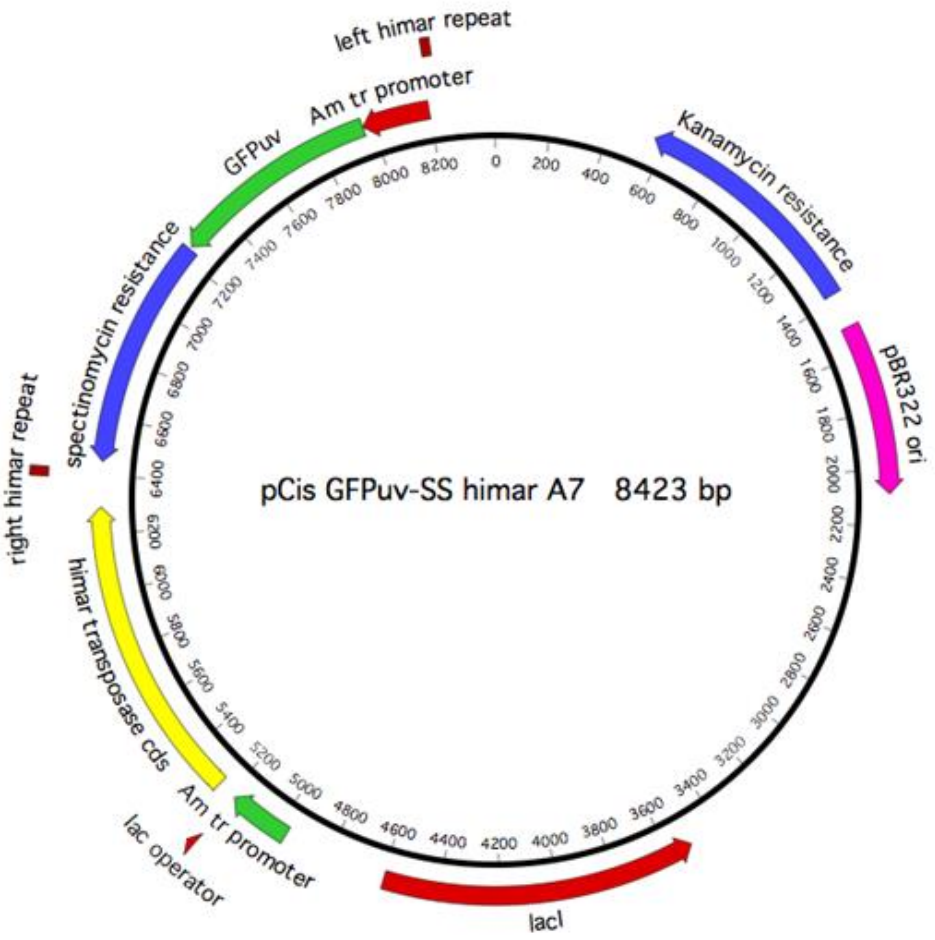


Figure 3.2: **pCis GFPUV –SS A7 –Himar plasmid**

This plasmid is essentially the same of pCis cherry –SS A7 –himar, except that the mCherry gene coding sequence is replaced with GFP-UV coding sequence.

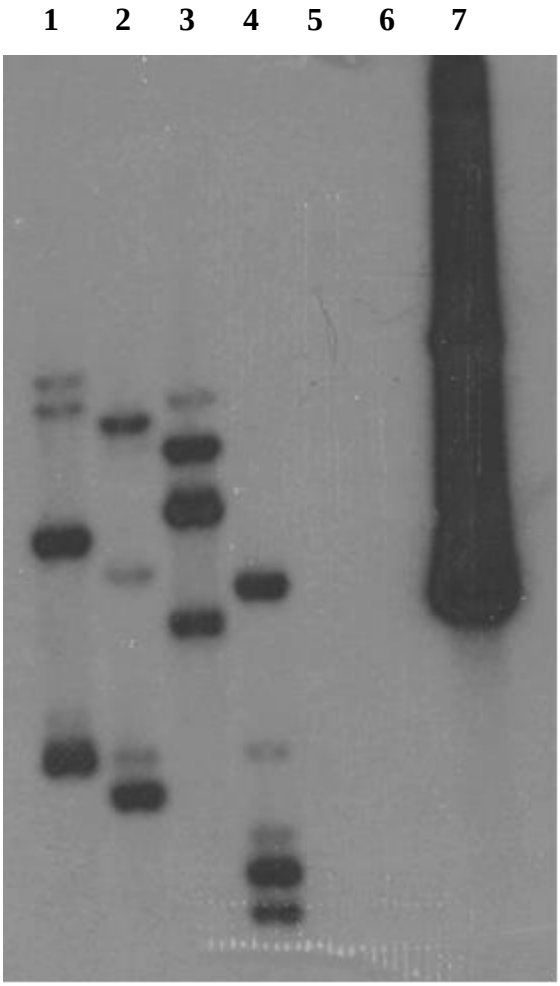


Figure 3.3 : Southern blot experiment

Genomic DNAs were isolated from *E. chaffeensis* transposon mutants prepared using the pCis GFPuv–SS A7 or pCis mCherry–SS A7–himar plasmids and digested with BgII (lanes 1 and 3) or Hind III (lanes 2 and 4). The DNAs were resolved on a 0.9 % agarose gel, transferred on to a nylon membrane and probed with a ^{32}P labeled spectinomycin resistance gene specific probe. Lanes 1 and 2, genomic DNA isolated from mCherry mutants; lanes 3 and 4, genomic DNA from GFPuv mutants; lanes 5, wild-type DNA, lane 6 blank lane, and lane 7, PCR product used for probe synthesis to serve as the positive control.

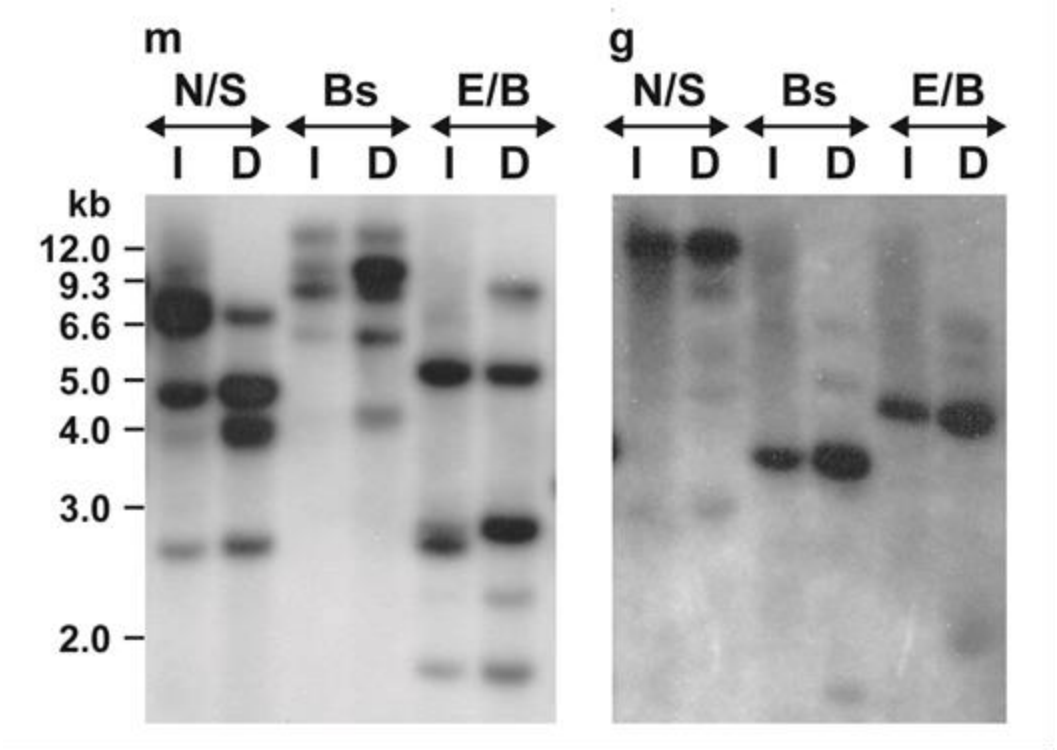


Figure 3.4: Southern blot experiment

Genomic DNA from Himar1 transposon E. chaffeensis mutants was assessed by DNA blot analysis using a spectinomycin resistance gene (aad) probe following digestions with different restriction enzymes (N, NdeI; S, Spe I; Bs, Bsrg I; E, EcoR V; B, Bgl II and Hind III). *E. chaffeensis* genomic DNA was recovered from the mutant organisms of mCherry plasmid transformed (m) or GFPuv plasmid transformed organisms (g) grown in ISE6 tick cells (I) or DH82 macrophages (D). It was digested with different restriction enzymes, resolved on a 0.9 % agarose gel and transferred on to a nylon membrane. Then they were hybridized using a ^{32}P labeled aad gene specific probe.

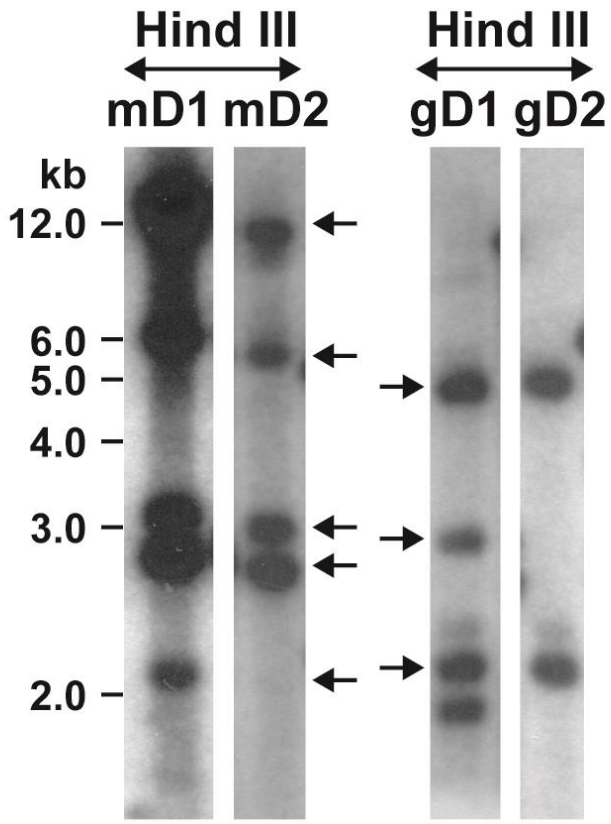


Figure 3.5: Southern blot experiment

Genomic DNAs from the mCherry and GFPuv (1st experiment) mutants propagated in DH82 cultures was also assessed at two different randomly selected harvest times (separated about 1 month apart) to evaluate the stability of the transformants over time. The lanes mD1 and mD2 represent two different days when mCherry mutants were harvested, and gD1 and gD2 represent different harvest dates for GFPuv mutants.

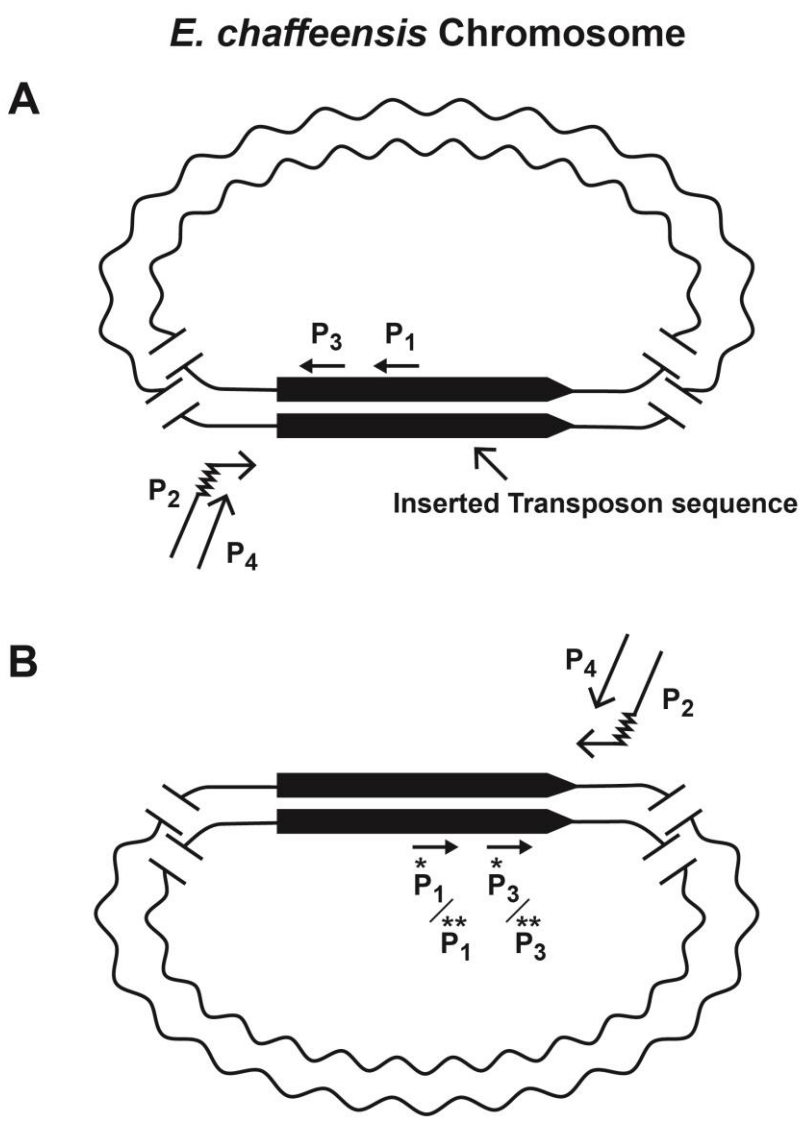


Figure 3.6: Cartoon Illustrating the ST-PCR

The sets of primers P1/P1*/P1** and P3/P3*/P3** are primers specific to inserted transposon sequence and P2 and P4 are the random primers. The random primer P2 contains three segments; the 5' end defined sequence, the middle segment containing ten random nucleotides and the 3' end contained GATAT. P4 has just the 5' end defined sequence.

P 2 (RRG 1260) - 5' GGCCACGCGTCGACTAGTACN $_{(10)}$ GATAT 3'

P4 (RRG 1261) - 5' GGCCACGCGTCGACTAGTAC 3'

$P_1/P_1*/P_1**$ - RRG 1259, RRG 1194, RRG1264

$P_3/P_3*/P_3**$ - RRG 1258, RRG 1194, RRG 1202

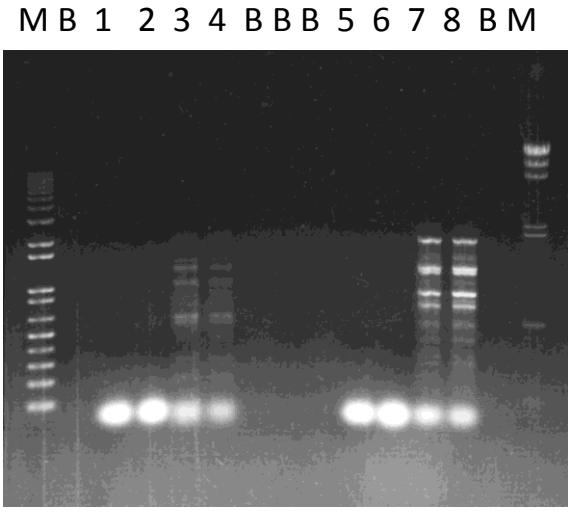


Figure 3.7 First round ST-PCR

First round ST-PCR with random primer RRG 1260 (P_2) paired with RRG 1259 (lanes 2 and 6), or RRG 1264 (lanes 3, 4, 7 and 8) (P_1, P1* respectively). Lanes 1 and 5 are negative controls; lanes 2, 3 and 4 are the ST-PCR products using GFPuv transposon mutants genomic DNA as the template; Lanes 6, 7 and 8 are the ST-PCR products from the mCherry transposon mutants genomic DNA. Lanes M at the left and right ends contained 1 KB + DNA ladder and lambda DNA digested with Hind III restriction enzyme, respectively. PCR products were resolved on a 0.9% agarose gel.

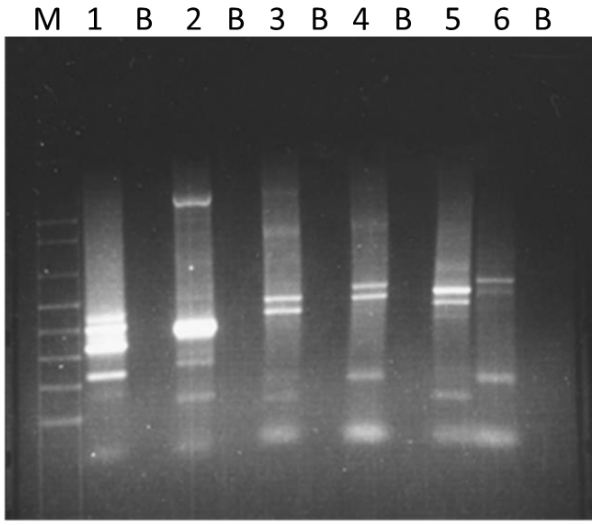


Figure 3.8: Second round ST-PCR

Second round ST-PCR performed with P_4 (RRG 1261) and with (P_3, P_3*, P_3**) RRG 1258 (Lane 3 and 5), or RRG1194 (lanes 4 and 6), or RRG1202 (lanes 1 and 2) . Lanes 1, 2and 3 are GFP; lanes 4, 5 and 6 are mCherry derived DNA. B, blank lane, Lane M is 1kb + DNA ladder. PCR products were resolved on a 0.9 % agarose gel at 90 volts/ 2hours.

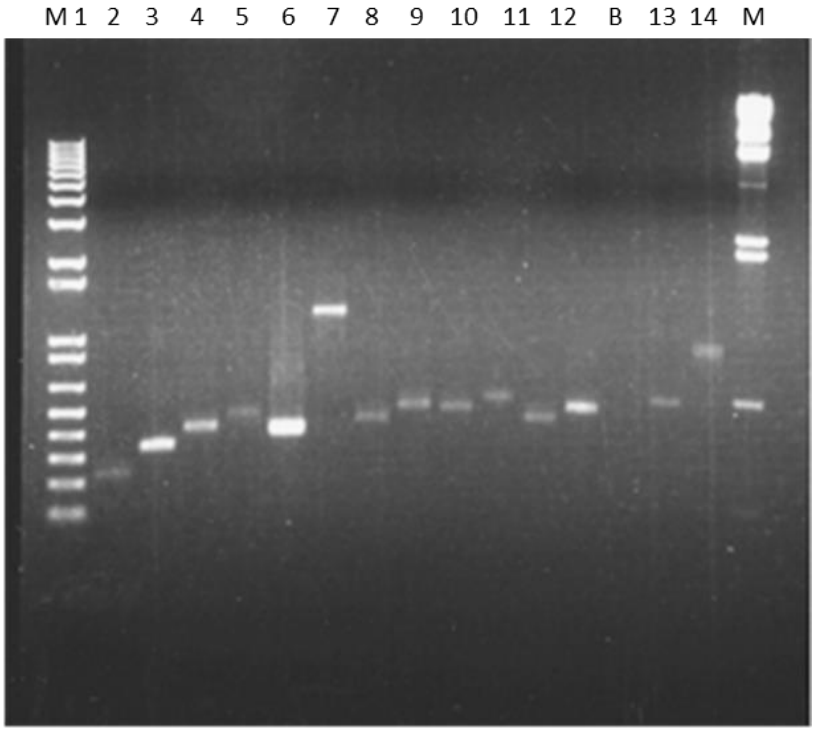


Figure 3.9: Gel isolated ST-PCR products

The ST-PCR DNA bands from the gel (figure 3.7) were isolated and purified. All the samples were resolved on 0.9 % agarose gel; Lanes 1 through 14 represent the purified DNA from different bands isolates from the resolved gel. Lanes M at the left and right ends contained 1 KB + DNA ladder and lambda DNA digested with Hind III restriction enzyme, respectively.

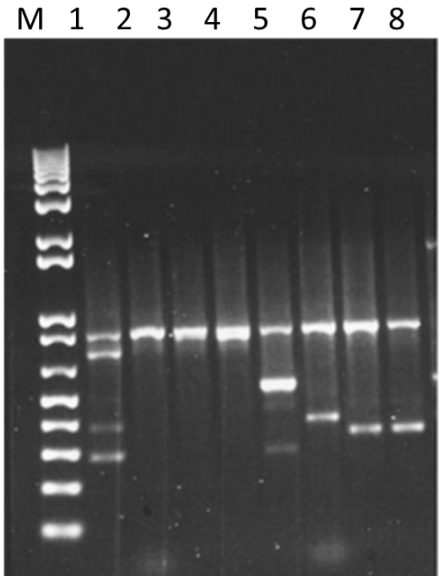


Figure 3. 10: 2nd ST-PCR experiment- first round PCR

First round ST-PCR with Random primer RRG 1260 (P_2) paired with RRG 1194 (lane 3 and 7) or, RRG 1259 (lanes 1, 2, 5 and 6), or RRG 1264 (lanes 4 and 8) ($P^*/P^*_1/P^{**}_1$). Lanes 1, 2, 3 and, 4 are GFP; lanes 5, 6, 7 and, 8 are mCherry. Lane M is L- 1kb+ ladder. PCR products were resolved on a 0.9 % agarose gel.

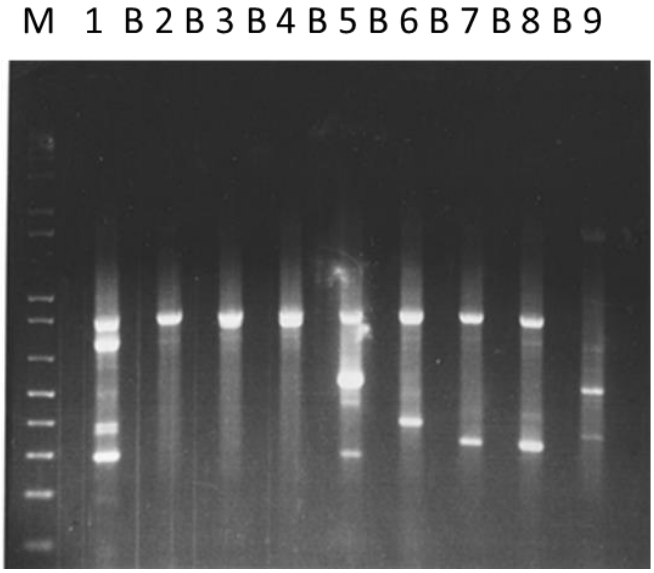


Figure 3.11: Second ST-PCR experiment- second round PCR

Second round ST PCR with P_4 (RRG 1261) and P_3/P^*_3/P^{**}_3; RRG 1258 (lanes 1, 2, 5, and 6) or RRG1194 (lanes 3 and 7) or RRG1202 (lanes 4, 8 and 9). Lanes 1, 2, 3, and 4 are GFP; lanes 5, 6, 7, 8 and 9 are mCherry. B, blank lane, lane M is 1kb + DNA ladder. Products resolved on a 0.9% gel at 100 volts /1.5 hours.

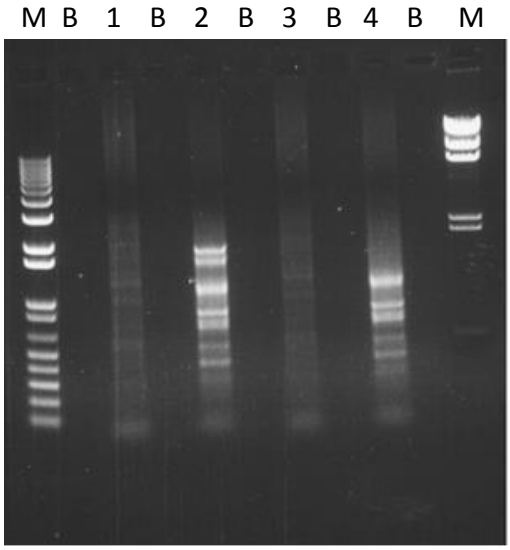


Figure 3.12: 3rd ST-PCR experiment –first round PCR

Third ST-PCR first round products resolved on 0.9 % agarose gel; 1st round ST-PCR with random primer RRG 1260 (P_2) paired with RRG 1194. Lanes 1 and 2 are GFP; lanes 3 and 4 are mCherry. Lanes M at the left and right ends contained 1 KB + DNA ladder and lambda DNA digested with Hind III restriction enzyme, respectively.

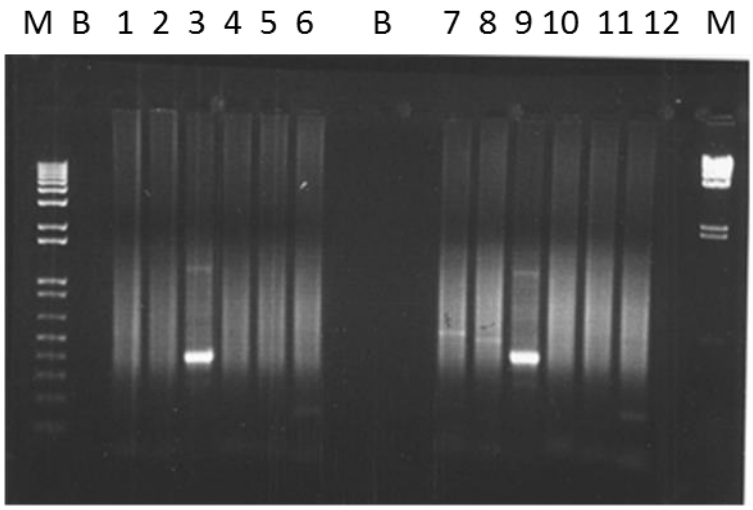


Figure 3. 13: 3rd ST-PCR experiment – second round PCR

Third ST-PCR with random primers RRG 1260 paired with RRG 1258. Second round products resolved on 0.9% agarose gel; 2 dilutions of the1st round products were used as templates for 2nd round. Lanes 1, 2, 3 – GFP (dilution 1- 1:5 dilution); lanes 4, 5, 6 - mCherry (dilution 1). Lanes 7, 8, 9 - GFP (dilution 2); lanes 10, 11, 13 – mCherry (dilution 2 is 1:10 dilution). Lanes M at the left and right ends contained 1 KB + DNA ladder and lambda DNA digested with Hind III restriction enzyme, respectively

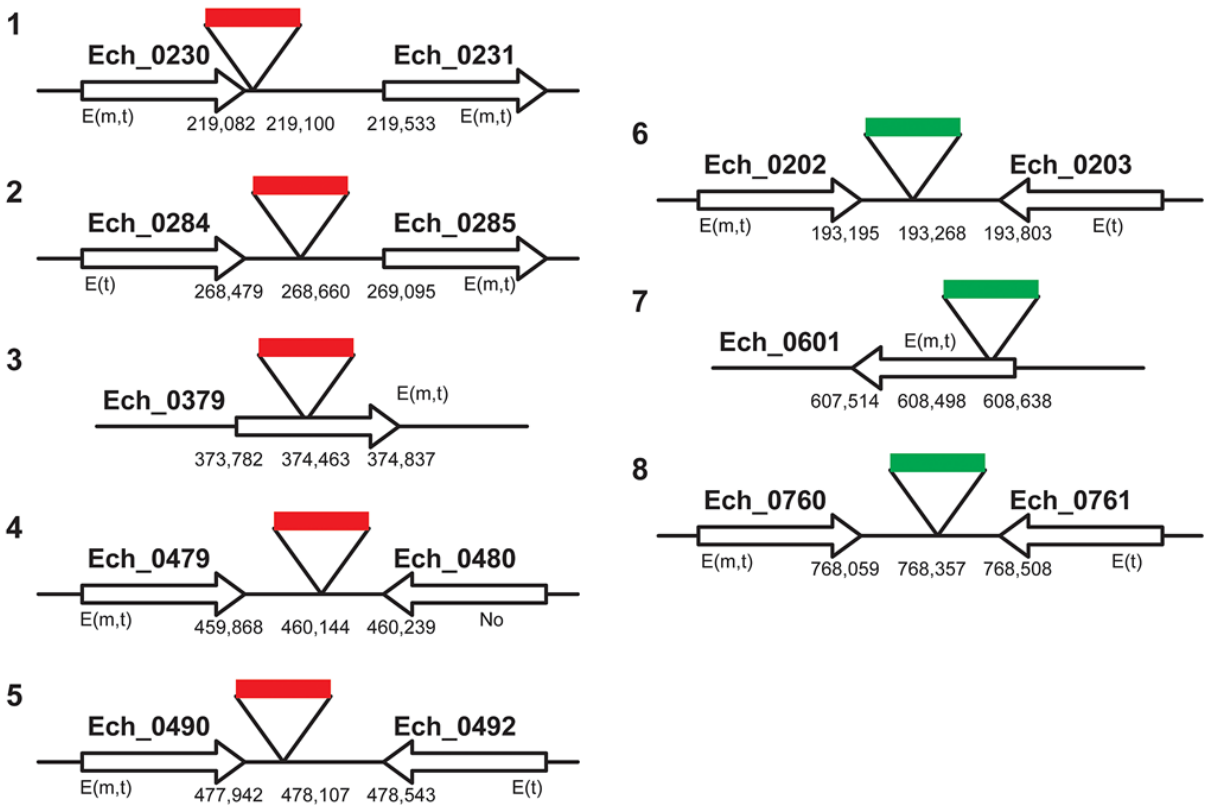
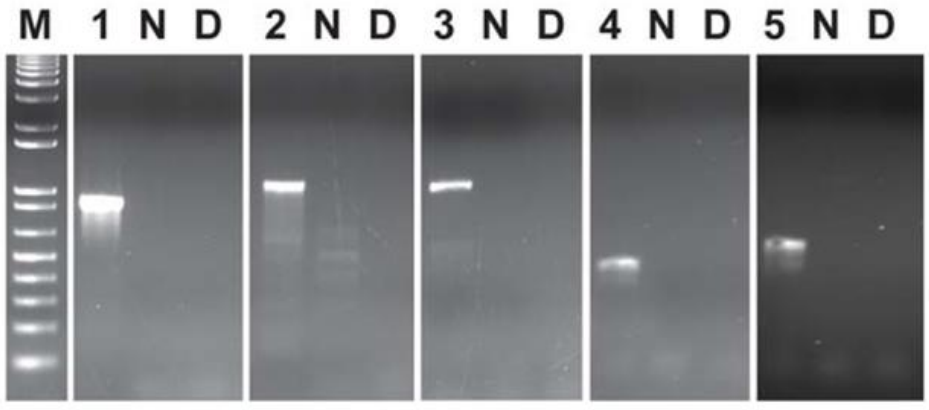


Figure 3.14: Mapped transposon insertion sites

A cartoon illustration of the E. chaffeensis genomic locations mapped for the transposon mutants. E. chaffeensis genomic DNA from three independent transformations with mCherry and GFPuv Himar1 transposon plasmids was used to determine the integration locations by inverse PCRs and ST-PCRs followed by DNA sequence analysis. Genomic locations of the insertion sites and the genes at or near the insertions, as per the whole genome data (GenBank # CP000236.1), were presented. The insertions in mCherry transformants are shown as solid red bars, and insertions in GFPuv transformants are depicted as solid green bars



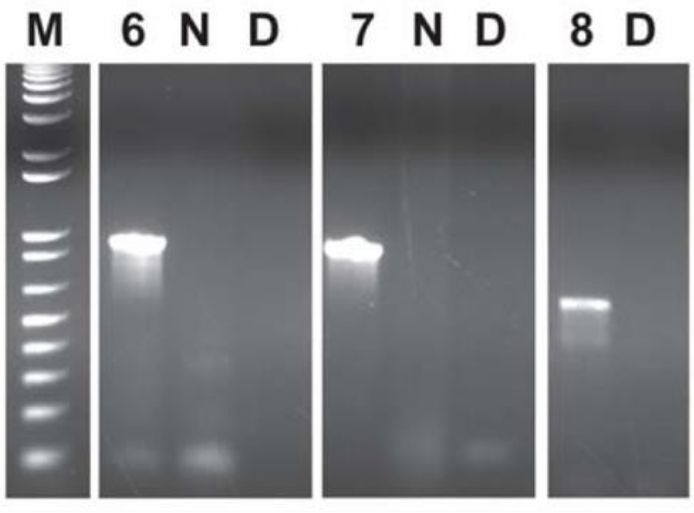


Figure 3.15: Validation of transposon insertions in the *E. chaffeensis* genome.
The insertion sites in the E. chaffeensis genome were verified by PCR with primers designed to bind to the genomic region upstream of the insertion sites (forward primer) and to the inserted DNA (spectinomycin resistance gene) (reverse primer). Product sizes for all 8 insertions are different and the predicted size amplicons were observed only in PCRs containing the mutant genomic DNAs as the templates. Lanes 1- 8 different mutants. N, no template control; D, wild type E. chaffeensis DNA used as the template; M, 1 kb+ DNA molecular weight marker.

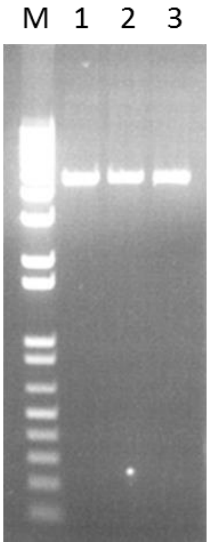


Figure 3.16: Restriction digestion of Himar1-UV-SS plasmid (single digestion)

The transposon plasmid pHIMAR1-UV-SS isolated from overnight *E.coli* cultures was digested with restriction enzyme EcoRI to make it linear; The products were resolved on 0.9% agarose gel AT 108 VOLTS / 2 hours ; the band size is 4.5 kb. M- 1kb plus DNA molecular weight marker.

Lanes 1, 2 and 3 – plasmid DNA isolated from three clones.

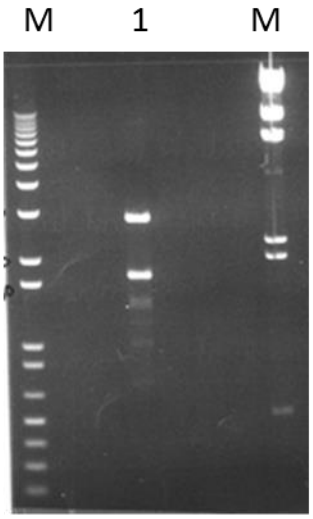


Figure 3.17: Restriction digestion of Himar1-UV-SS plasmid (double digestion)
Plasmid pHIMAR1-UV-SS digested with two restriction enzymes with ECORI and XbaI.Products were resoled on 0.9% agarose gel at 120 volts/2hours. The lower band is around 1755 bp and the upper band is 2804 bp. M (far left) is 1kb + ladder, M (far right) is Hind III digested lambda DNA marker**.**

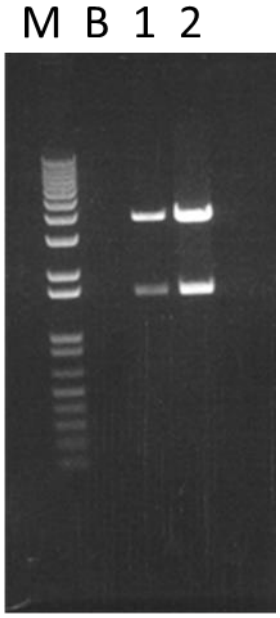


Figure 3.18: Restriction digestion of ML.cat224 plasmid

Transposon plasmid ML.cat 224 containing Echtuf promoter and chloramphenicol acetyl transferase resistance gene was digested with restriction enzyme SpeI to release the 1.6 kb segment containing Echtuf promoter and chloramphenicol acetyl transferase resistance gene. The product was resolved on 0.9 % gel. Lanes 1 and 2 are plasmid isolated from 2 clones. Lane M is 1kb plus DNA ladder.

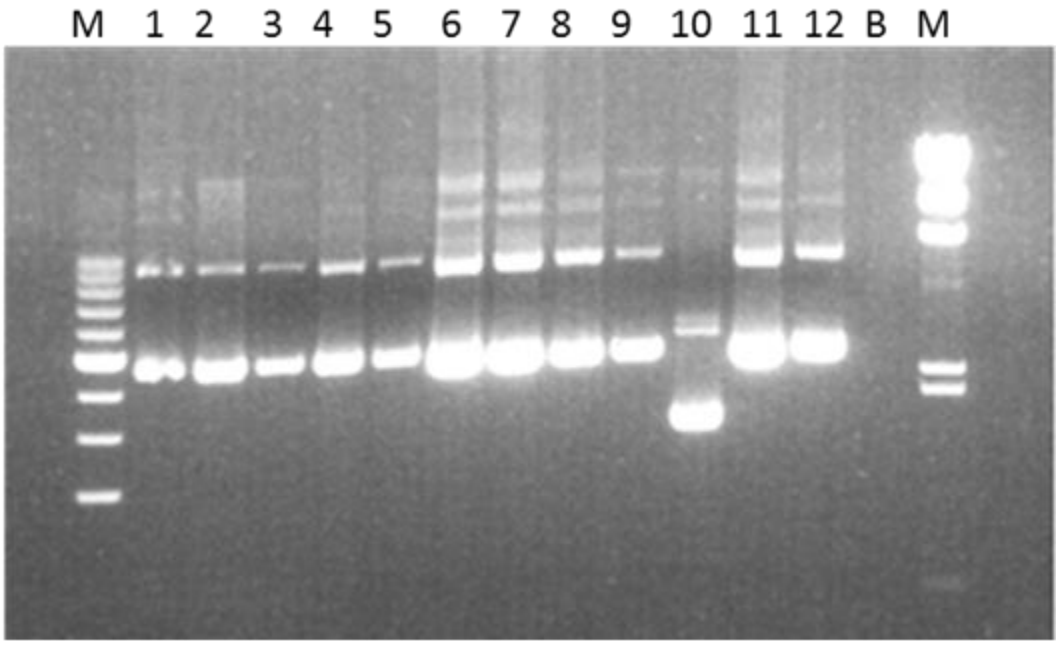


Figure 3.19: pHimar UV-SS cat 224 plasmid selection

1.6kb gene segment containing Echtuf promoter and chloramphenicol acetyl transferase resistance gene form the Ml.cat 224 plasmid was ligated in to the Himar- uvss plasmid vector to create pHimar UV-SS cat 224.

Lanes 1- 12 - Himar-uvss.Ml.cat 224 plasmids isolated from several clones; uncut and resolved on 0.9 % gel; Lanes M at the left and right ends contained 1 KB + DNA ladder and lambda DNA digested with Hind III restriction enzyme, respectively.

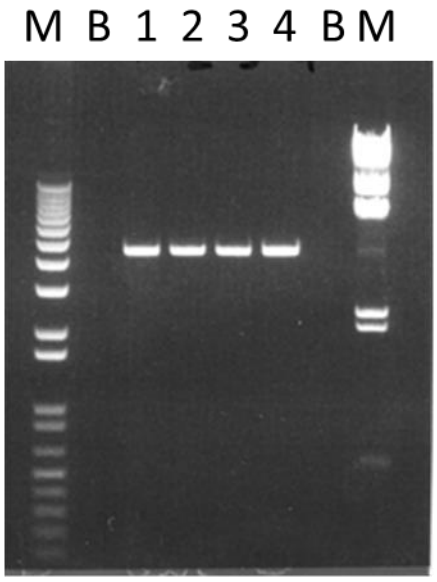


Figure 3.20: pHimar UV-SS cat 224 plasmid verification by single restriction enzyme digestion
The new transposon plasmid Himar-uvss.Ml.cat 224 plasmid digested with single restriction enzyme Hind III to give 4.5 kb size fragment. Lanes 1, 2, 3 and 4 are four different samples; M (far left) is 1 Kb + DNA ladder, M (far right) is Hind III digested Lambda DNA. Products were resolved on 0.9% gel at 120 volts/ 1.5 hours.

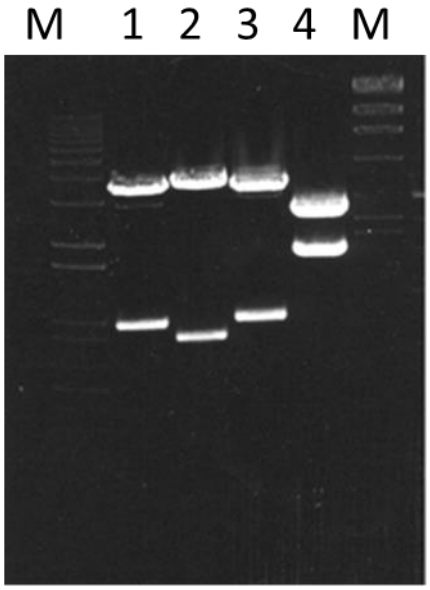


Figure 3.21: pHimar UV-SS cat 224 plasmid verification by single and double restriction enzyme digestion

The transposon plasmid Himar-uvss.Ml.cat 224 plasmid digested with single and double restriction enzymes. Lane 1 is XhoI digestion, lane 2 is SalI digestion, lane 3 is NcoI digestion and lane 4 is Spe I and Hind III double digestion. M (far left) is 1 Kb + DNA ladder M (far right) is Hind III digested Lambda DNA. Products were resolved on 0.9% gel at 120 volts/ 1.5 hours

Chapter 4

Discussion

E. chaffeensis is an emerging pathogen causing disease in both people (HME) and various vertebrate animals which include dog, coyote, goat and deer [1, 8]. HME was first discovered in 1987 and subsequently reported in many parts of the USA and several other countries from western hemisphere, Europe, and Asia [89-93]. Despite the significant health concern, much remains to be understood about how *E. chaffeensis* causes pathogenesis and how it successfully survives in tick and vertebrate hosts. Recent studies suggest that the pathogen alters its protein expression in dual host environments and the protein expression differences are also shown to be influencing the host response and the pathogen persistence [94]. An important goal of the research for our group is to understand the molecular basis of the *E. chaffeensis* pathogenesis and also to understand the host-specific differences in gene expression of the pathogen. One of the means to investigate the importance of host-specific gene expression differences is by employing mutational analysis methods. In support if this, we conducted experiments to create transposon mutations in *E. chaffeensis.* Specifically, Himar transposon mutagenesis was performed which resulted in creating mutations in multiple genomic locations. In this study, I utilized a rapid and efficient *in vitro* amplification method to locate the genomic sites where the transposon insertions were introduced. The ST-PCR method used in this study aided in identifying and validating 8 genomic sites in *E. chaffeensis* genome. Experiments described in my research are also important in further verifying the insertion mutation locations by conducting insertion-specific PCRs and sequencing analysis. Together, the data described aided in the identification of mutations within the coding regions of two genes, Ech_0379 and Ech_0601 and one at immediate downstream to the coding region of a

gene, Ech_0230. These three mutations caused the loss of gene activity (research results of our group). Five mutations were also identified at intergenic sites of *E. chaffeensis* chromosome and these mutations had no impact of the expression genes located at both upstream and downstream from the insertion sites. The RNA analysis was performed and published as part of the study conducted by our research team [88]. Animal infection studies also conducted by our research team further demonstrated that two of the three gene inactivation mutations at Ech_0230 and Ech_0379 caused attenuated growth of the mutant organisms *in vivo* when assessed in white-tailed deer. Together, these data demonstrate the value of the mutagenesis in identifying genes critical for the pathogen's growth *in vivo*. Importantly, defining the location of mutations within the genome of the organism is a critical step in initiating the detailed investigations of understanding the importance of specific gene expression to the pathogen. Moreover, the methods established in this study will lead the way for additional studies to create mutations at numerous genomic locations in *E. chaffeensis* and to determine which genes are critical for the pathogen's *in vivo* growth and in causing pathogenesis. These studies are now actively pursued by our research team

As part of the MS research, I also modified the existing transposon plasmids (prepared earlier for use in *Anaplasma* species mutational analysis) for use in improving the efficiency of transposon mutagenesis in *E. chaffeensis*. In particular, a transposon insertion plasmid was by replacing the *Anaplasma* species promoter with a constitutively expressed *E. chaffeensis* gene promoter, *Ech-tuf*. Likewise, the antibiotic resistance cassette conferring resistance to spectinomycin was replaced with chloramphenicol resistance gene, CAT. This modified construct, although not tested as

part of my current research project, is likely valuable for future experiments in improving the transposon efficiency in *E. chaffeensis*. This is an important goal for our research because our first set of transposon mutagenesis experiments resulted in a very limited number of mutants. (Eight mutants were reported in the current study.) Although this is an impressive achievement for the *Ehrlichia* field as this is the first mutational analysis study, high efficiency mutagenesis is needed to map the functions of large numbers of genes present in the genome of *E. chaffeensis*. Therefore, the modifications to Himar mutagenesis plasmids are important in optimizing the mutational experiments for the generation of mutant library containing large numbers of mutants.

In summary, the research conducted as part of my MS graduate education led to the standardization and the utilization of genome mapping methods needed for locating transposon insertion sites efficiently. This research also is a stepping point for improving the efficiency of transposon mutagenesis in *E. chaffeensis*.

Chapter 5 - References

1. Lockhart, J.M., et al., *Temporal association of Amblyomma americanum with the presence of Ehrlichia chaffeensis reactive antibodies in white-tailed deer.* J.Wildl.Dis., 1995. **31**: p. 119-124.
2. Dumler, J.S., et al., *Isolation and characteriztion of a new strain of Ehrlichia chaffeensis from a patient with near-fatal monocytic ehrlichiosis.* Journal of Clinical Microbiology, 1995. **33**: p. 1704-1711.
3. Dumler, J.S., et al., *Human granulocytic anaplasmosis and Anaplasma phagocytophilum.* Emerging Infectious Diseases, 2005. **11**(12): p. 1828-1834.
4. Paddock, C.D., et al., *Brief Report: Fatal seronegative Ehrlichiosis in a patient with HIV infection.* New England Journal of Medicine, 1993. **329**: p. 1164-1167.
5. Dawson, J.E., et al., *White-tailed deer as a potential reservoir of Ehrlichia spp.* J.Wildl.Dis., 1994. **30**: p. 162-168.
6. Dugan, V.G., et al., *Detection of Ehrlichia spp. in raccoons (Procyon lotor) from Georgia.* Vector.Borne.Zoonotic.Dis., 2005. **5**(2): p. 162-171.
7. Maeda, K., et al., *Human infection with Ehrlichia canis, a leukocytic rickettsia.* New England Journal of Medicine, 1987. **316**(14): p. 853-856.
8. Dawson, J.E., et al., *Isolation and characterization of an Ehrlichia sp. from a patient diagnosed with human ehrlichiosis.* Journal of Clinical Microbiology, 1991. **29**(12): p. 2741-2745.
9. Fishbein, D.B., J.E. Dawson, and L.E. Robinson, *Human ehrlichiosis in the United States, 1985 to 1990.* Annals of Internal Medicine, 1994. **120**: p. 736-743.

10. Paddock, C.D. and J.E. Childs, *Ehrlichia chaffeensis: a prototypical emerging pathogen.* Clin.Microbiol.Rev.2003.Jan.;16.(1.):37.-64. **16**: p. 37-64.

11. Olano, J.P., et al., *Clinical manifestations, epidemiology, and laboratory diagnosis of human monocytotropic ehrlichiosis in a commercial laboratory setting.* Clin.Diagn.Lab Immunol., 2003. **10**(5): p. 891-896.

12. Bakken, J.S., et al., *Serological evidence of human granulocytic ehrlichiosis in Norway.* Eur.J.Clin.Microbiol.Infect.Dis., 1996. **15**: p. 829-832.

13. Whitlock, J.E., et al., *Prevalence of Ehrlichia chaffeensis (Rickettsiales: Rickettsiaceae) in Amblyomma americanum (Acari: Ixodidae) from the Georgia coast and Barrier Islands.* J.Med.Entomol., 2000. **37**(2): p. 276-280.

14. McNabb, S.J., et al., *Summary of notifiable diseases--United States, 2006.* MMWR Morb.Mortal.Wkly.Rep., 2008. **55**: p. 1-92.

15. Bakken, J.S., et al., *Human granulocytic ehrlichiosis in the upper Midwest United States. A new species emerging?* JAMA, 1994. **272**(3): p. 212-218.

16. Rikihisa, Y., *Anaplasma phagocytophilum and Ehrlichia chaffeensis: subversive manipulators of host cells.* Nature Reviews Microbiology, 2010. **8**: p. 328-339.

17. Parola, P., B. Davoust, and D. Raoult, *Tick- and flea-borne rickettsial emerging zoonoses.* Vet.Res., 2005. **36**: p. 469-492.

18. Dumler, J.S., et al., *Ehrlichioses in humans: epidemiology, clinical presentation, diagnosis, and treatment.* Clinical.Infectious.Diseases., 2007. **45 Suppl 1**(16): p. S45-S51.

19. Paddock, C.D., et al., *Infections with Ehrlichia chaffeensis and Ehrlichia ewingii in persons coinfected with human immunodeficiency virus.* Clinical.Infectious.Diseases., 2001. **33**: p. 1586-1594.

20. Dumler, J.S. and J.S. Bakken, *Ehrlichial diseases of humans: emerging tick-borne infections.* Clinical.Infectious.Diseases., 1995. **20**: p. 1102-1110.

21. Gardner, S.L., et al., *National surveillance for the human ehrlichioses in the United States, 1997-2001, and proposed methods for evaluation of data quality.* Annals of the New York Academy of Sciences, 2003. **990:80-9.**: p. 80-89.

22. Anderson, B.E., et al., *Ehrlichia chaffeensis, a new species associated with human ehrlichiosis.* Journal of Clinical Microbiology, 1991. **29**(12): p. 2838-2842.

23. Chapman, A.S., et al., *Diagnosis and management of tickborne rickettsial diseases: Rocky Mountain spotted fever, ehrlichioses, and anaplasmosis--United States: a practical guide for physicians and other health-care and public health professionals.* MMWR Recomm.Rep., 2006. **55**: p. 1-27.

24. Uematsu, C., et al., *Multiplex polymerase chain reaction (PCR) with color-tagged module-shuffling primers for comparing gene expression levels in various cells.* Nucleic.Acids.Res., 2001. **29**: p. E84.

25. Stone, J.H., et al., *Human monocytic ehrlichiosis.* JAMA, 2004. **292**(18): p. 2263-2270.

26. Brouqui, P., M.L. Birg, and D. Raoult, *Cytopathic Effect, Plaque Formation, and Lysis of Ehrlichia chaffeensis Grown on Continuous Cell Lines.* Infection and Immunity, 1994. **62**: p. 405-411.

27. Asanovich, K.M., et al., *Antigenic diversity of granulocytic Ehrlichia isolates from humans in Wisconsin and New York and a horse in California.* Journal of Infectious Diseases, 1997. **176**: p. 1029-1034.

28. Dunning Hotopp, J.C., et al., *Comparative Genomics of Emerging Human Ehrlichiosis Agents.* PLoS.Genet., 2006. **2**: p. e21.

29. Lin, M. and Y. Rikihisa, *Ehrlichia chaffeensis and Anaplasma phagocytophilum lack genes for lipid A biosynthesis and incorporate cholesterol for their survival.* Infection and Immunity, 2003. **71**: p. 5324-5331.

30. Hotopp, J.C., et al., *Comparative genomics of emerging human ehrlichiosis agents.* PLoS.Genet., 2006. **2**: p. e21.

31. Singu, V., et al., *Unique Macrophage and Tick Cell-specific Protein Expression from the p28/p30 Omp Multigene Locus in Ehrlichia Species.* Cell Microbiol., 2006. **8**: p. 1475-1487.

32. Seo, G.M.G.R.R., *Proteome analysis of E. chaffeensis in different host cell backgrounds*, 2007.

33. Seo, G.M., et al., *Total, membrane, and immunogenic proteomes of macrophage- and tick cell-derived Ehrlichia chaffeensis evaluated by LC-MS/MS and MALDI-TOF methods.* Infection and Immunity, 2008. **76**: p. 4823-4832(PMID: 18710870).

34. Lin, Q., et al., *Analysis of sequences and loci of p44 homologs expressed by Anaplasma phagocytophila in acutely infected patients.* J Clin.Microbiol., 2002. **40**(8): p. 2981-2988.

35. Yu, X., et al., *Characterization of the complete transcriptionally active Ehrlichia chaffeensis 28 kDa outer membrane protein multigene family.* Gene, 2000. **248**: p. 59-68.

36. Yu, X.J., et al., *Molecular cloning and characterization of the 120-kilodalton protein gene of Ehrlichia canis and application of the recombinant 120-kilodalton protein for serodiagnosis of canine ehrlichiosis.* J Clin.Microbiol., 2000. **38**: p. 369-374.

37. Ohashi, N., Y. Rikihisa, and A. Unver, *Analysis of Transcriptionally Active Gene Clusters of Major Outer Membrane Protein Multigene Family in Ehrlichia canis and E. chaffeensis.* Infection and Immunity, 2001. **69**: p. 2083-2091.

38. Reddy, G.R., et al., *Molecular characterization of a 28 kDa surface antigen gene family of the tribe Ehrlichiae.* Biochemical and Biophysical Research Communications, 1998. **247**: p. 636-643.

39. Cheng, C., C.D. Paddock, and G.R. Reddy, *Molecular Heterogeneity of Ehrlichia chaffeensis Isolates Determined by Sequence Analysis of the 28-Kilodalton Outer Membrane Protein Genes and Other Regions of the Genome.* Infection and Immunity, 2003. **71**: p. 187-195.

40. Ohasi, N., et al., *Immunodominant major outer membrane proteins of Ehrlichia chaffeensis are encoded by a polymorphic multigene family.* Infection and Immunity, 1998. **66**: p. 132-139.

41. Chao, C.C., D. Chelius, and D.L.C.W. Zhang T, *Proteome analysis of Madrid E strain of Rickettsia prowazekii* Proteomics, 2004. **4**: p. 1280-1292.

42. van, H.H., et al., *Characterization of a major outer membrane protein multigene family in Ehrlichia ruminantium.* Gene., 2004. **330:159-68.**: p. 159-168.

43. Winslow, G.M., E. Yager, and J.S. Li, *Mechanisms of humoral immunity during Ehrlichia chaffeensis infection.* Annals of the New York Academy of Sciences, 2003. **990:435-43.**: p. 435-443.

44. Felek, S., et al., *Sensitive Detection of Ehrlichia chaffeensis in Cell Culture, Blood, and Tick Specimens by Reverse Transcription-PCR.* Journal of Clinical Microbiology, 2001. **39**: p. 460-463.

45. Wakeel, A., et al., *Ehrlichia chaffeensis Tandem Repeat Proteins are Related to the RTX Family of Toxins and Secreted by the Type 1 Secretion System.* Frontiers in Cellular and Infection Microbiology, 2011. **1**.

46. Little, S.E., et al., *Natural coinfection of a white-tailed deer (Odocoileus virginianus) population with three Ehrlichia spp.* Journal of Parasitology, 1998. **84**: p. 897-901.

47. Rikihisa, Y., H. Stills, and G. Zimmerman, *Isolation and continuous culture of Neorickettsia helminthoeca in a macrophage cell line.* Journal of Clinical Microbiology, 1991. **29**: p. 1928-1933.

48. Rikihisa, Y., et al., *Ultrastructural and antigenic characterization of a granulocytic ehrlichiosis agent directly isolated and stably cultivated from a patient in New York state.* Journal of Infectious Diseases, 1997. **175**: p. 210-213.

49. Rikihisa, Y., *Clinical and biological aspects of infection caused by Ehrlichia chaffeensis.* Microbes.Infect., 1999. **1**: p. 367-376.

50. Yu, D.H., et al., *Ehrlichia chaffeensis infection in dogs in South Korea.* Vector.Borne.Zoonotic.Dis., 2008. **8**(3): p. 355-358.

51. Ewing, S.A., et al., *Experimental transmission of Ehrlichia chaffeensis (Rickettsiales: Ehrlichieae) among White-tailed deer by Amblyomma americanum (Acari: Ixodidae).* J.Med.Entomol., 1995. **32**: p. 368-374.

52. Unver, A., et al., *Western and dot blotting analyses of Ehrlichia chaffeensis indirect fluorescent-antibody assay-positive and -negative human sera by using native and*

recombinant E. chaffeensis and E. canis antigens. Journal of Clinical Microbiology, 1999. **37**: p. 3888-3895.

53. Telford, S.R., et al., *Perpetuation of the agent of human granulocytic ehrlichiosis in a deer tick-rodent cycle.* Proc.Natl.Acad.Sci.U.S.A., 1996. **93**: p. 6209-6214.

54. Rikihisa, Y., *Mechanisms to create a safe haven by members of the family Anaplasmataceae.* Annals of the New York Academy of Sciences, 2003. **990**: p. 548-555.

55. Kiristioglu, I., et al., *Total parenteral nutrition-associated changes in mouse intestinal intraepithelial lymphocytes.* Dig.Dis.Sci., 2002. **47**: p. 1147-1157.

56. Webster, P., et al., *The agent of Human Granulocytic Ehrlichiosis resides in an endosomal compartment.* Journal of Clincal Investigation, 1998. **101**: p. 1932-1941.

57. Huang, H., et al., *Proteomic analysis of and immune responses to Ehrlichia chaffeensis lipoproteins.* Infection and Immunity, 2008. **76**: p. 3405-3414(PMID: 18490460).

58. Dedonder, S.E., et al., *Transmission Electron Microscopy Reveals Distinct Macrophage- and Tick Cell-Specific Morphological Stages of <italic>Ehrlichia chaffeensis</italic>.* PLoS ONE, 2012. **7**(5): p. e36749.

59. Popov, V.L., et al., *Ultrastructural differentiation of the genogroups in the genus Ehrlichia.* Journal of Medical Microbiology, 1998. **47**: p. 235-251.

60. Popov, V.L., et al., *Ultrastructural variation of cultured Ehrlichia chaffeensis.* Journal of Medical Microbiology, 1995. **43**: p. 411-421.

61. Munderloh, U.G., et al., *Invasion and intracellular development of the human granulocytic ehrlichiosis agent in tick cell culture.* Journal of Clinical Microbiology, 1999. **37**(8): p. 2518-2524.

62. Zhang, J.z., et al., *The developmental cycle of Ehrlichia chaffeensis in vertebrate cells.* Cellular Microbiology, 2007. **9**: p. 610-618.

63. Munderloh, U.G., et al., *Isolation of an Anaplasma sp. organism from white-tailed deer by tick cell culture.* Journal of Clinical Microbiology, 2003. **41**: p. 4328-4335.

64. Troese, M.J., et al., *Differential expression and glycosylation of anaplasma phagocytophilum major surface protein 2 paralogs during cultivation in sialyl Lewis x-deficient host cells.* Infect Immun., 2009. **77**(5): p. 1746-1756.

65. Cheng, Z., X. Wang, and Y. Rikihisa, *Regulation of type IV secretion apparatus genes during Ehrlichia chaffeensis intracellular development by a previously unidentified protein.* Journal of Bacteriology, 2008. **190**(6): p. 2096-2105.

66. Wolf, K., E. Fischer, and T. Hackstadt, *Ultrastructural analysis of developmental events in Chlamydia pneumoniae-infected cells.* Infection and Immunity, 2000. **68**(4): p. 2379-2385.

67. Winslow, G.M., et al., *Antibody-mediated elimination of the obligate intracellular bacterial pathogen Ehrlichia chaffeensis during active infection.* Infection and Immunity, 2000. **68**: p. 2187-2195.

68. Ganta, R.R., et al., *Persistent Ehrlichia chaffeensis infection occurs in the absence of functional major histocompatibility complex class II genes.* Infection and Immunity, 2002. **70**: p. 380-388.

69. Ganta, R.R., et al., *Delayed clearance of Ehrlichia chaffeensis infection in CD4+ T-cell knockout mice.* Infection and Immunity, 2004. **72**: p. 159-167.

70. Dinarello, C.A., *Proinflammatory and anti-inflammatory cytokines as mediators in the pathogenesis of septic shock.* Chest, 1997. **112**: p. 321S-329S.

71. Bitsaktsis, C., J. Huntington, and G. Winslow, *Production of IFN-gamma by CD4 T cells is essential for resolving ehrlichia infection.* Journal of Immunology, 2004. **172**: p. 6894-6901.

72. Rudel, T., et al., *Interaction of two variable proteins (PilE and PilC) required for pilus-mediated adherence of Neisseria gonorrhoeae to human epithelial cells.* Mol.Microbiol., 1992. **6**: p. 3439-3450.

73. Giron, J.A., et al., *The flagella of enteropathogenic Escherichia coli mediate adherence to epithelial cells.* Mol.Microbiol., 2002. **44**: p. 361-379.

74. Ganta, R.R., et al., *Differential clearance and immune responses to tick cell-derived versus macrophage culture-derived Ehrlichia chaffeensis in mice.* Infection and Immunity, 2007. **75**: p. 135-145(PMID: 17060466).

75. Singu, V., et al., *Ehrlichia chaffeensis expresses macrophage- and tick cell-specific 28-kilodalton outer membrane proteins.* Infection and Immunity, 2005. **73**: p. 79-87.

76. Ganta, R.R., et al., *Molecular characterization of Ehrlichia interactions with tick cells and macrophages.* Front Biosci., 2009. **14**: p. 3259-3273 (PMID: 19273271).

77. Sirigireddy, K.R., et al., *Gene expression profiles of Ehrlichia chaffeensis originating from vertebrate and tick cell environments*, 2009.

78. Ijdo, J.W., A. Muller, and A. Carlson, *Transposan-based mutagenesis of Anaplasma phagocytophilum.* 4th International conference on Rickettsiae and Rickettsial diseases, 2005.

79. Wood, D.O. and A.F. Azad, *Genetic manipulation of rickettsiae: a preview.* Infection and Immunity, 2000. **68**: p. 6091-6093.

80. Hughes, A.L. and J.O. French, *Homologous recombination and the pattern of nucleotide substitution in Ehrlichia ruminantium.* Gene., 2006. **.**

81. Felsheim, R.F., et al., *Transformation of Anaplasma phagocytophilum.* BMC.Biotechnol., 2006. **6**: p. 42.

82. Felsheim, R.F., et al., *An unexpected Anaplasma marginale transformant.* 21st Meeting of the American Soceity for Rickettsiology, 2007.

83. Rachek, L.I., et al., *Transformation of Rickettsia prowazekii to erythromycin resistance encoded by the Escherichia coli ereB gene.* J Bacteriol., 2000. **182**: p. 3289-3291.

84. Felsheim, R.F., et al., *Transformation of Anaplasma marginale.* Veterinary Parasitology, 2010. **167**(2–4): p. 167-174.

85. KRISTIN T. CHUN, H.J.E., MARK R. KELLY AND mARK G. GOEBL, *Rapid Amplification of Uncharacterized Transposon tagged DNA sequences from Genomic DNA* Yeast 1996. **13**: p. 233-240.

86. Shevchuk, O., et al., *InFiRe a novel computational method for the identification of insertion sites in transposon mutagenized bacterial genomes.* Bioinformatics, 2012. **28**(3): p. 306-310.

87. Cheng, C. and R.R. Ganta, *Laboratory Maintenance of Ehrlichia chaffeensis and Ehrlichia canis and Recovery of Organisms for Molecular Biology and Proteomics Studies*, in *Current Protocols in Microbiology*. 2005, John Wiley & Sons, Inc.

88. Cheng, C., et al., *Targeted and Random Mutagenesis of <italic>Ehrlichia chaffeensis</italic> for the Identification of Genes Required for <italic>In vivo</italic> Infection.* PLoS Pathog, 2013. **9**(2): p. e1003171.

89. Bakken, J.S., et al., *Clinical and laboratory characteristics of human granulocytic ehrlichiosis.* JAMA, 1996. **275**: p. 199-205.

90. Bakken, J.S., et al., *Seroprevalence of human granulocytic ehrlichiosis among permanent residents of northwestern Wisconsin [In Process Citation]*. Clinical.Infectious.Diseases., 1998. **27**: p. 1491-1496.

91. Dumler, J.S. and J.S. Bakken, *Human ehrlichioses: newly recognized infections transmitted by ticks.* Annu.Rev.Med., 1998. **49**(19): p. 201-213.

92. Standaert, S.M., et al., *Primary isolation of Ehrlichia chaffeensis from patients with febrile illnesses: clinical and molecular characteristics.* Journal of Infectious Diseases, 2000. **181**: p. 1082-1088.

93. Olano, J.P. and D.H. Walker, *Human ehrlichioses.* Med.Clin.North Am., 2002. **86**: p. 375-392.

94. Cotter, P.A. and V.J. DiRita, *Bacterial virulence gene regulation: an evolutionary perspective.* Annual Review of Microbiology, 2000. **54**: p. 519-565.